

The end of political gridlock, at last?

US President-elect Bill Clinton is putting pragmatism ahead of ideology in his first major appointments of people with some sympathy for a national science policy.

WITHIN the past several days, US President-elect Bill Clinton, with Vice-President and future technology czar Al Gore at his side, has named people to cabinet and other high-level positions in his administration that, taken together, give clues about how the Clinton White House will handle science and technology in its drive to improve the US economy.

First, by choosing Texas Senator Lloyd Bentsen as Treasury Secretary, Clinton has eschewed the anti-business stance of liberal Democrats, perhaps taking a cue from former presidential-candidate Paul Tsongas of Massachusetts who urged the party to recognize that the government cannot distribute wealth that has not been created. Bentsen is sympathetic to giving tax advantages to business, such as tax credits for research and development. Two other senior economic appointees are on record as supporting active government intervention in ways that could affect science and technology.

Industrial policy

Political economist Robert Reich of Harvard University, a friend of Clinton's from his days as a Rhodes Scholar who has been named Secretary of Labor, favours active policies to revive US industry. And Berkeley economist Laura Tyson, who is likely to become chairman of the White House Council of Economic Advisors (CEA), goes so far as to support the idea of a formal industrial policy (something the United States has previously shied away from).

The appointment of a strong environmentalist as head of the Environmental Protection Agency (EPA) also signals a change in emphasis for that agency. Carol M. Browner, who formerly worked for Gore in the Senate, and is, at present, Secretary of the Florida State Department of Environmental Regulation, has a reputation as an activist who is able to compromise with industry to get things done.

Even in his candidate for Secretary of Health and Human Services, University of Wisconsin chancellor Donna Shalala, Clinton has chosen someone sympathetic to the idea of a coordinated national policy for science and technology. Shalala, known as an educator and expert on urban problems, also has experience with the National Institutes of Health (NIH) where she is a member of the director's advisory committee.

At the top of the Clinton team (which is still incomplete) sits Gore whose knowledge of (and interest in) science and technology exceeds not only that of previous vice-presidents but also that of the president-elect himself. During the

campaign, Clinton promised to shift research and development resources from the Department of Defense (DOD) to the civilian sector. Clinton said he is interested in a civilian 'ideas' agency similar to DOD's Defense Advanced Research Projects Agency (DARPA) and pledged to provide funds to foster collaboration between academic science and industry. It is expected that the job of carrying out these promises will fall to Gore, who is eager, willing and, most important, qualified to do the job.

Picture still incomplete

Many second-tier positions within the new administration need to be filled before the picture of the Clinton team is complete but some things can be said about what the president-elect should do. It is not clear how influential the White House Office of Science and Technology Policy, headed by the president's science adviser, will be in a government with so many other strong players (Gore may well overshadow the science adviser). The best way to ensure some strength in the position would be to appoint (for the first time) a science adviser whose expertise is in the biological rather than physical sciences, thereby complementing Gore's expertise in environment and technology.

Similarly, in a government intent on fostering the idea that science and technology can and should be bent to national economic and competitive needs, Clinton would do well to select as heads of the NIH and the National Science Foundation scientists who will be strong defenders of basic research, initiated at the laboratory bench and not in federal committee rooms.

It is tempting to wax enthusiastic about a new administration with obvious appreciation of the importance of science and technology in the United States' national and international life — tempting but much too soon. But it is possible to speculate that the new crew, taken together, portend a government that will try to reach compromise on complex and controversial issues, in part by providing more coordination from the top.

This is both good and potentially dangerous. A policy that places the practical uses of science ahead of the continuing need to support unfettered research is a danger the Clinton administration must avoid. But if a consensus among top government officials on general science and technology policy can lead to the end of the gridlock that has paralysed Washington for the past several years, the Clinton administration will have accomplished a great deal. □



MRC attacks government proposal to sell off research institutes

London. Tensions over the fate of Britain's five research councils broke into the open last week when the head of one, the Medical Research Council, publicly criticized a proposal that the council should sell off its research units and institutes, and restrict itself to distributing research funds. The proposal includes selling such bodies as the MRC's Laboratory of Molecular Biology (LMB) in Cambridge and the National Institute for Medical Research in Mill Hill.

Dai Rees, the secretary to the council, called the idea a "disaster" and said that research institutes are "at the heart of the MRC's strategy". Privatizing the institutes would encourage them to focus on the short-term needs of the market rather than the long-term needs of society, he said, a change that was "deeply and fundamentally misguided".

The proposals by the Advisory Committee on Science and Technology (ACOST), if adopted, would represent one of the biggest shifts in the MRC's 80-year history and could mark the end of the council in its present form. Although the directors of many of the research institutes would welcome greater flexibility in the way they operate, they do not want to sever direct links with the MRC.

ACOST's proposals are contained in a report drawn up by a subcommittee headed by Sir Robin Nicholson, a science adviser to the former prime minister, Margaret Thatcher. They were put together at the request of William Waldegrave, the Chancellor of the Duchy of Lancaster, whose white paper (policy document) on the reorganization of British science is due early next year.

One proposal that has attracted controversy is to replace the five research councils with two funding bodies. One, known as the Council for the Advancement of Scientific Knowledge (CASK), would assume responsibility for 'curiosity-driven' research; the other, a Council for Mission-Oriented Research (COMR), would handle all research linked to clear social objectives.

The second proposal is to privatize almost all publicly owned laboratories, including those currently run by the research councils. They would either be sold off to companies or universities, or encouraged to become independent contract-research organizations.

ACOST admits that special steps would be needed to maintain the quality of research carried out in the privatized institutes and that certain functions should be kept in the public domain. Councils such as the MRC would become purely funding agencies, operating as "programme boards" of the

new COMR unless they chose to give up any policy-making role and become merely managers of research institutes.

Rees believes that the ACOST proposals run counter to the MRC's own experience. Close contact between basic and clinical research is essential at all levels of research management, he says, citing as an example monoclonal antibodies, which were discovered in the mid-1970s in the MRC's Cambridge laboratories and have since generated a world-wide market valued at £200 million a year.

Rees also denies that the so-called "protected access" to funding bodies enjoyed by its research institutes has lessened scientific quality. He points out that the MRC's recent decision to close a number of its institutes — most recently being its Epidemiology Unit in South Wales — demonstrates that scientific quality remains essential for continued funding.

Aaron Klug, director of the Cambridge laboratory, is equally critical of the recommendations in the ACOST report, saying that "privatization would not do us any good at all". He calls the proposal "a totally retrograde step, because it sets up a false divide between the so-called 'purchasers'

and 'providers' of research".

Nicholson says that fears of what might happen to research institutes such as the LMB under the ACOST proposals are being exaggerated. "We do not want to destroy the long-term relationship between a research council and an institute and have suggested that this relationship be maintained through long-term rolling contracts", he says. "In particular, ACOST does not think there is anything wrong with the LMB. But we do think that there are things wrong with some of the other research institutes, and can trace these back to their ownership by government departments."

It is not known how much support the proposals have within ACOST. At least two committee members, Sir Michael Atiyah, president of the Royal Society, and Sir David Phillips, the chairman of the Advisory Board for the Research Councils, have publicly expressed their disagreement with some of its conclusions. At the same time, ACOST's conclusion that research institutes should be sold off and operated as private agencies is in line with government policy that has already, for example, been applied to the former United Kingdom Atomic Energy Authority.

David Dickson

Politics alleged in Indian fraud

New Delhi. Political interference in investigations of scientific fraud is once again at the centre of a storm in the university town of Chandigarh, northern India. The latest controversy involves Chandigarh's Postgraduate Institute for Medical Research, where the institute's ethics committee has upheld allegations of wrongdoing against one of its staff, dermatologist Bhusnan Kumar. But whereas Kumar is being allowed to remain, the head of his department, Surrinder Kaur, who brought the fraud to light, has resigned and is leaving the institute.

At issue is a paper coauthored by Kumar in the *Indian Journal of Dermatology, Venereology and Leprology* claiming that the herpes zoster virus had been cultured in egg yolk. Kaur doubted the claim and alleged that the cultures "were not even attempted" and that the data had been fabricated. A nine-member ethics committee reached the same conclusion after a six-month investigation, stating that "the scientific data reported [were] not based on facts". Kumar confessed to the editor of the journal that "the culture results reported cannot be substantiated" and that

he felt sorry for the "glaring error".

Nonetheless, the institute's governing body has decided against sacking Kumar, saying that he has apologized for the "lapse" and citing his potential as a researcher. Kumar is also a relative of an influential politician who is a member of the governing body. Kaur says that the decision "was an administrative whitewash" and that she quit to focus attention on "blatant political interference in academic affairs". She has asked India's health minister, the chairman of the governing body, to review the case.

The new furore is an echo of similar events at the Panjab University, also in Chandigarh, where palaeontologist Vishwa J. Gupta has clung to his position despite a massive amount of evidence implicating him in the falsification of fossil records in the Himalayas (see *Nature* 355, 579; 1992). Observers say that the two cases demonstrate the erosion of ethical and academic values at Chandigarh and the futility of exposing misconduct when those accused have powerful political friends.

K.S. Jayaraman

Among the rice fields, Japan tests the international waters

Tokyo. Japan's first big test of how well it handles the internationalization of science is unfolding in the middle of a rice-farming region 100 km northeast of Tokyo. Over the coming months, dozens of researchers from Europe, the United States and Russia will arrive at the Naka fusion research establishment to operate Japan's contribution to the International Thermonuclear Experimental Reactor (ITER) project. Naka is one of three co-centres — the others are in Garching, Germany and San Diego, California — for the six-year, \$1 billion engineering design phase of ITER (see *Nature* 358, 269; 1992).

Naka was chosen by Japan's Science and Technology Agency (STA) because Japan's JT-60 Tokamak is located there and there are many scientists and engineers at the fusion establishment with expertise in the design and building of superconducting magnets, one of the key components of ITER that the Naka group will be responsible for designing. And at present there is plenty of room for foreigners. Michel Huguet, deputy director of the design phase for the European Communities and head of the Naka co-centre; Britain's Barry Green; and R.O. Thome of the United States are the only three non-Japanese researchers now working at Naka. Over the next two years, however, they will be joined by more than 40 scientists from Europe, the United States and Russia as the engineering design phase gets into full swing.

That influx will overwhelm the Japanese contingent, making Naka the first example

in Japan of a truly international research centre dominated and directed by non-Japanese. Japan's lack of experience in operating such an international organization is already beginning to show.

One immediate problem is the lack of an international school for the children of the non-Japanese researchers. The closest established schools are 100 km away in Tokyo, although Tsukuba science city, about halfway between Tokyo and Naka, does have a fledgling international elementary school for a handful of young children (see below).

STA and the Japan Atomic Energy Research Institute, which runs the Naka establishment, have hurriedly begun building a small international school at Naka to be operated as a branch of St Mary's International School of Tokyo. However, the school, scheduled to open in a few months, will enrol children only up to 15 years of age. Older children must attend international schools in Tokyo and stay with Japanese families if their parents choose to live in Naka.

One of the non-Japanese researchers already at Naka has a teenage daughter.



Above: An aerial view of the Naka Fusion Research Establishment.



Left: Opening ceremonies for the Naka centre were held in October.

Rather than have his daughter live separately from the family, the researcher has decided to live in Tokyo and to commute to Naka.

The Japanese are doing their best to accommodate the foreign researchers. Twenty single-family houses are being built in Naka, and unmarried scientists will be housed in brand-new apartments that are very spacious by Japanese (and even Western) standards. Furthermore, the foreign scientists will not have to pay the expensive fees for their children to attend international school.

Another problem concerns support personnel. Under the agreement signed in July for the engineering design phase, Japan has to provide several 'designers' at Naka. But there has been disagreement about the requisite skills of such designers. Japan at first offered computer-assisted design (CAD) operators who can handle data only in Japanese and who have no ability to design a thermonuclear reactor. The other ITER partners want designers with backgrounds in mechanical engineering and an ability to communicate in English.

Shin Aoyama, director for fusion energy at STA, says CAD designers with those skills are very hard to find in Japan. In the West, suitable designers can be found for moderate salaries among people without a university-level education. But in Japan such designers are scarce and command high salaries. This week, when the ITER Council meets in Moscow, Japan is expected to give a firm commitment to provide the necessary level of designers.

So things at Naka are moving in the right direction, says one organizer of ITER. But he says that, despite their goodwill, the Japanese do not fully seem to realize that the Naka centre must be organized along internationally established lines and not on Japanese terms because the majority of researchers will be non-Japanese.

David Swinbanks

Tsukuba school paves the way

Tokyo. After years of lobbying by foreign scientists, Tsukuba science city northeast of Tokyo at last has an international school to educate some of the children of the city's large community of foreign scientists.

Tsukuba International School, which opened in September, is only a small start — a private elementary school with about a dozen children aged 6 to 11 and one British teacher. Nevertheless, it is a significant step forward for Japan, where government science policy makers repeatedly call for "internationalization" but seldom provide or even encourage the necessary local infrastructure.

There are about 150 foreign school-age children in Tsukuba and the number is growing each year. Until now, these children have had to attend local Japanese schools, which pose problems for older children because the classes are in Japanese.

The school was started by a small local group of foreign and Japanese volunteers with money from private companies and one government organization. A local cram school provided a classroom at token rent.

The effort has taken three years and one large hurdle remains. The school has not existed long enough to be granted tax-free status as a *gakko hojin* (school foundation) from the local prefectural education board under the ministry of education; as a result, companies are reluctant to contribute to it because their donations are taxed.

Another problem is that scientists at public research institutes cannot afford its fees. The school has introduced some scholarships but, again, funds are limited because it lacks *gakko-hojin* status.

D.S.

DARPA's small satellite programme shrinks

Washington. Recent cuts in spending for small US satellite technology leave this still-developing field without one of its main sources of funding, highlighting the precarious foothold that 'smallsats' have in the overall US space programme.

Congress, in its 1993 defence appropriations bill, deleted \$30.4 million for three programmes requested by the Defense Advanced Research Projects Agency (DARPA). Two would have led to tests in space of new, miniaturized communications and remote sensing technologies, while the third would have produced a common satellite 'bus' to provide cheaper and quicker flight opportunities for small orbiting payloads.

Congress agreed to spend \$18 million for other DARPA space activities, however, including the first flight next spring of the Taurus mobile launch system for small payloads.

More than any other government agency, DARPA has acted as an incubator for the fledgling small satellite and small launch industries. It has financed development of two new small launchers — the Pegasus aeroplane-launched rocket and the new Taurus vehicle — both built by the Orbital Sciences Corporation (OSC) in Fairfax,

Virginia. Other DARPA research has been directed at reducing the size, weight and cost of components to be used in future military satellites.

Small satellites, say advocates, hold out the hope of cheaper and more frequent flights that would accommodate more people. They would give scientists better access to Earth orbit for experiments in astronomy, space physics, materials science and related fields.

Critics of the three programmes — the Advanced Satellite Technology and EHF Communications (ASTEC), the Collaboration on Advanced Multispectral Earth Observation (CAMEO) and the Advanced Technology Standard Satellite Bus (ATSSB) — say they serve no military needs that aren't already being met by existing programming. And with Congress eager to trim defence spending, the estimated \$147 million lifetime cost of the three projects was too tempting a target to ignore.

John Pike, a space programme analyst with the Federation of American Scientists, says that the dilemma for DARPA's space technology programme is that "either this stuff has direct mission applications, in which case it is competing with existing programmes with powerful constituencies, or it

does not have direct mission applications, in which case it's an awfully expensive way to advance generic technology".

But Jill Stern of the International Small Satellite Organization, which supports the development of smallsats, thinks that DARPA should be engaged in pure research. "The point is to develop technology", she says.

Nongovernment researchers for the most part have ignored Pegasus flights, which cost between \$11 million and \$15 million. With few exceptions, the money to buy OSC's small rockets has come from federal agencies such as DARPA, the Strategic Defense Initiative (SDI) and the National Aeronautics and Space Administration (NASA). The dearth of outside customers for such flights raises the question of how much longer the government should operate such incubators if there is no commercial interest in them.

DARPA's role in developing smallsat technology may depend on how President-elect Bill Clinton redefines the mission of the 34-year-old military research agency. During the presidential campaign, Clinton suggested a civilian equivalent of DARPA for commercial technology development. Others have proposed similar concepts, ranging from a quasi-governmental Civilian Technology Corporation to a restructured DARPA for both military and civilian high technology.

A recent analysis by the Rand Corporation criticizes both ideas, however, by arguing that any federal agency is ill-equipped to deal with the constant changes in commercial technology. In addition, even a new structure would eventually lose its flexibility and develop the same institutional biases that plague existing agencies.

With DARPA's fate hanging in the balance, supporters of small satellites hope that the three programmes will be revived in the agency's budget for fiscal year 1994 beginning on 1 October 1993. If not, they must place their hopes for funding with the SDI programme, which faces its own problems in a Clinton administration, or with NASA, which has shown a growing interest in small satellites.

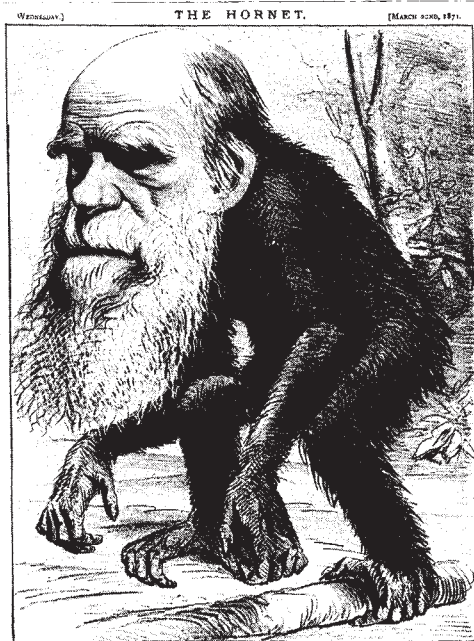
Tony Reichhardt

Corrections

■ The European Societies of Pharmacology were wrongly referred to as the European Federation of Pharmacology Societies in a recent News article about homeopathic products (*Nature* 359, 469; 1992).

■ The picture of the Moon that appeared in last week's News section (*Nature* 360, 501; 1992) was taken by the Galileo spacecraft not during this month's trip around the Earth but during its initial flyby two years ago.

Darwin material sells briskly at Sotheby's



Evolution went under the hammer in London last week, when Sotheby's, the auctioneers, sold a collection of more than 400 books, pamphlets and manuscripts relating to Charles Darwin and his colleagues. The collection had been built up over the past 25 years by a Californian scientific book collector, Jeremy Norman. A single manuscript page referring to Creation and natural selection from Charles Darwin's *On the Origin of Species* was sold for £9,350 — more than twice the pre-auction estimate — and a page from *The Descent of Man* went for £4,400; a first edition of the latter, estimated at between £2,000 and £2,500, fetched £4,400. Included in the auction was set of anti-Darwinian pamphlets (at right), published in the 1870s, which sold for £235.

David Dickson

Revamped Delbrück centre enters new era

Heidelberg. Bringing East Berlin's Max Delbrück Centre for Molecular Medicine up to western standards has been a hard and painful process. But Detlev Ganten, its director, believes that the worst is over

summoned to Moscow by Joseph Stalin to dissect Lenin's brain as a way of unravelling the 'anatomy of genius'. (Although Vogt found nothing more unusual than a few enlarged cells in the hippocampus, Stalin heralded the observation as the source of all genius.)

The institute survived the war and was retained as a research centre — expanding into three institutes — by the East German National Academy of Science. In 1989, after reunification, the research centre once again demonstrated its power of survival by passing muster with the Bonn-based AGF, which funds and administers national research centres. Only one of its three institutes, the Max Stelbourg, has been redeveloped, along with cardiovascular and cancer clinics that together have 300 beds.

As with all East German institutes, however, the Max Stelbourg was heavily over-staffed. And Ganten, who became director in September 1991, has had the unhappy task of shedding more than half of the 1,600 employees. Recognizing the need to staunch the outflow of talent, Ganten retained as many researchers as possible on short-term contracts and arranged for 250 others to receive special government 'rescue' grants

that allowed them to continue working without actually being employed. He also encouraged the formation of individual pharmaceutical companies (25 now exist) by staff from the applied pharmaceuticals department in the old research centre.

The streamlined Max Delbrück centre now plans to support 20 permanent professorships or group leaders, as well as 350 employees in experimental research, 112 of whom are research scientists, and 400 people in the two clinics. Ganten's goal is to give permanent positions to 20–30 per cent of the staff, a level appropriate to the competitive field of molecular medicine. Equally important to Ganten is restoring the centre's links to the universities. Both clinics are incorporated into the Free University of Berlin, and professors of experimental medicine will be affiliated with either the Humboldt or the Free universities.

Despite the project's success in western terms, Ganten remains sensitive to the depression that pervades eastern Germany. He says that his western compatriots are still seen as "brutal in their efficiency, colonizing their country and forcing a system on them that has caused hardship for so many".

Nonetheless, Ganten is proud of what he has done. He abandoned a comfortable life at the University of Heidelberg for a gamble that, so far, appears to be paying off.

Alison Abbott



The reopened Delbrück centre

and that the revamped centre, which reopened earlier this month, stands a good chance of living up to its illustrious past.

A biological institute on the Max Delbrück site was founded in the 1930s by the Kaiser Wilhelm Society, the forerunner of the Max Planck Society. The institute, which specialized in brain research, attracted the best of German scientists as well as occasional notoriety, as when neurologist Oskar Vogt was

\$100 million pledged to support ex-Soviet science

Washington. Plans by a US philanthropist to spend \$100 million on science in the former Soviet Union during the next several years must somehow find a way to overcome the logistical nightmare of finding and funding the right scientists.

George Soros, a financier who made a billion dollars earlier this year by betting on the devaluation of the British pound, announced last week that his new International Science Foundation for the Former Soviet Union will dispense its first grants early next year. But foundation officials admit that the target countries are just starting to create a system to handle such largesse.

The gift dwarfs existing efforts to help ex-Soviet scientists. The United States, European Communities, Canada and Japan have pledged \$70 million to help former weapons scientists in Russia and another \$12 million for those in Ukraine, but after nearly a year the first grants have yet to be awarded. The American Physical Society, with help from Soros, the Sloan Foundation and the National Science Foundation, has amassed \$700,000 to distribute directly to

needy scientists, avoiding government bureaucracy. And the Howard Hughes Medical Institute has just announced a five-year programme to spend \$14 million to bolster research in the former Soviet Union.

According to preliminary plans, the foundation will first devise a system for distributing the bulk of the money. About \$6 million will be spent on 'emergency projects', with half allocated to salaries and the rest to libraries and a telecommunications system. Scientists from every former Soviet republic can qualify by filling out a brief, simplified grant application or by having already undergone peer review in the West.

In the second phase of the project, the foundation plans to solicit Western-style applications and have them reviewed by committees of scientists from around the world. In mid-year, the foundation hopes to begin awarding 300–500 grants, averaging \$100,000, for salaries, supplies and equipment. The rest of the money will be spent on infrastructure. Soros plans to commit — if not spend — all the money by the end of 1994.

But the foundation faces obstacles that

have dissuaded other charitable groups. Because the republics lack a reliable banking system, foundation officials have proposed an electronic network for wiring money directly to researchers. But US regulations and high costs may make this plan unworkable. Nor will it be easy to ensure that the researchers who get funded are the most deserving. Many scientists who have never written a grant application will have only until mid-1993 to prepare one. "Those people who are quickest to learn how to write a proposal are not always the best", says Edward Shuryak, a physicist at the State University of New York, Stony Brook, who left the Soviet Union in 1989.

Other Soviet émigrés worry that the money will not be enough to make a permanent difference in an otherwise desperate economic situation. Even Alexander Goldfarb, the foundation's temporary director, has taken a wait-and-see attitude. "It's a huge experiment", Goldfarb says. "We'll see in six months whether it works or whether it just made a huge splash".

Traci Watson

British are helping to build Amazonian research station

São Paulo. In the Caxiuanã forest, 20 hours by boat west of Belém in the Amazonian state of Pará, a 126-year-old dream is becoming reality. With support from Britain's Overseas Development Administration, the Museu Paraense Emílio Goeldi in Belém is

from the British under the terms of an agreement signed several years ago for cooperative ventures to explore Brazil's rain forests. The station should be finished by May 1993 and will have accommodations and laboratory facilities for 60 scientists and technicians.

The foreign contribution shows that Brazil's research community can attract outside support, but the Goeldi museum is meeting the other 30 per cent of the cost on the promise, says Pedro L.B. Lisboa, the museum's vice-director for research, of important results. "Research programs and projects of high quality have always found a receptive audience among granting agencies both within Brazil and abroad", says Lisboa.

But he concedes that "the prospects are greater abroad".

The site in the Caxiuanã National Forest is pristine and sparsely populated, with varied vegetation. About 80 per cent of the 33,000 hectare site is tall, dry-land forest, the rest consisting of patches of floodable vegetation and savannah. The flora is among the richest in eastern Amazonia, according to Lisboa.

Ricardo Bonalume



The Goeldi Museum's research station in Amazonia.

building the tropical research station its directors have always wanted.

The Goeldi museum was begun by the nineteenth-century naturalist Domingos Soares Ferreira Penna, but named after Goeldi, a Swiss scientist who came to Brazil in the late 1800s and restored its flagging reputation. By way of historical justice, the forest research station will be named after Ferreira Penna. Seventy per cent of the US\$2.7 million construction cost comes

NEWS IN BRIEF

Munich. The heads of Germany's largest scientific research institutes and organizations last week asked their members to take an active stance against hostility towards foreigners.

The alliance, which represents nearly all German scientists, condemned racism and xenophobia and warned that "science, with its many international connections, is especially affected by foreigners' anxiety about Germany". The German research society, the Deutsches Forschungsgemeinschaft (DFG), has already reported that many of its foreign scientists and students now reject living in private accommodation, preferring student halls because of their perceived safety. A group of foreign scholars from the Humboldt Foundation, which promotes international scholastic exchange, were too frightened to leave their bus on a recent trip through cities in eastern Germany, including Rostock.

Institute officials emphasize that very few of their visitors have actually experienced hostility; instead, what they are likely to feel is fear. The statement by the alliance is backed by a national campaign using posters and other forms of advertising.

M.B.

Munich. The European Molecular Biology Conference, the intergovernmental body that finances the Heidelberg-based European Molecular Biology Organization (EMBO), has voted to include three new members: Turkey, Croatia and Portugal. The current membership is limited to 15 Western European countries, plus Israel.

The inclusion of three of Europe's poorest nations demonstrates the commitment of EMBO to the promotion of molecular biology throughout the continent. The political decision has also an element of generosity, as each member country pays as a fixed proportion of its gross national product. As a result, poor countries get out more than they put in.

A.A.

Vienna. Filippo Pandolfi, research minister for the European Communities (EC), last week announced the awarding of nearly 3,000 grants to scientists in Eastern European countries as part of an EC initiative to fund cooperative projects and exchange schemes between the old and the new democracies. Launched earlier this year as the largest programme of research cooperation between East and West, the initiative attracted an unexpectedly high number of applicants — 12,000. Most of the grants are for short fellowships under the Human Capital and Mobility scheme. Pandolfi said that the research ministry wants to repeat the programme next year but is dependent on a budget request discussed by EC ministers at their meeting last weekend in Edinburgh. This year's awards are worth Ecu55 million (US\$70 million).

A.A.

ESO offers aid to needy astronomers

Munich. The European Southern Observatory (ESO) is offering practical help to astronomers in eastern and central Europe whose working conditions have deteriorated during recent political upheavals.

ESO has allocated DM 500,000 (US\$310,000) of its DM120 million budget, which provides research facilities for its eight member states in Garching near Munich, for immediate aid to appropriate projects. The programme is primarily intended to help scientists to stay and work in their home institutes. While proposals will be judged on scientific merit, the fundamental needs of the beleaguered community are likely to mean that the grants will be used to ensure basic survival.

Some needs are easier to meet than others. In the last two years, for example, astronomers in Romania's three observatories have organized to assure their mutual

survival. But at the Bucharest Observatory 30 scientist must share a single personal computer, leaving some astronomers with only two hours a week of computing time. An ESO grant could remedy this situation by providing additional computers.

Other cases are more challenging. For example, the Byurakan Observatory in Armenia is cut off from access to the outside world and has no fuel and an unreliable supply of electricity. Despite the tremendous hardships, many scientists are trying to keep up some level of activity. However, the observatory is situated atop a mountain 1,500m high and must close during the winter.

ESO intends to invite some needy scientists to work for short periods of time at Garching. By acting as a clearing house, ESO also hopes to be able to channel support from other sources such as the European Communities.

Alison Abbott

Britain goes for collaboration in Europe

London. Whatever reservations the rest of the population may have about political and economic union with Europe, British scientists are becoming increasingly enthusiastic about collaborating with their colleagues across the Channel. A survey of more than 900 scientists involved in research projects sponsored by the Commission of the European Communities (EC) has shown that virtually all academic institutions in Britain receive money from Brussels.

An analysis of participation in the commission's framework programmes shows that Britain has the largest number of links of any of the 12 member-states of the European Communities and is the most frequent partner for seven of the 11 other states — including both France and Germany. In addition, four-fifths of those in the academic community receiving funds from Brussels say that it is making a significant or critical contribution to their research.

The survey was commissioned by Britain's Office of Science and Technology (OST) and the research directorate of the European Commission in Brussels and carried out by the Programme of Policy Research in Engineering, Science and Technology (PREST) at the University of Manchester and the Science Policy Research Unit (SPRU) at the University of Sussex.

The survey found that, overall, British universities receive about 5 per cent of their research funding through European programmes. The impact of this money is magnified, for example, by the fact that the Higher Education Funding Council gives extra funds to university departments that obtain research grants from Brussels.

"Our general conclusion is that, as a result of this funding, there has been a reorientation of the scientific community in this country to a point where British researchers now regard themselves as part of an emergent European scientific community", says Luke Georgiou of PREST, the leader of the research team.

The survey shows that 28 per cent of the projects funded from Brussels involve basic research; 42 per cent of the academic scientists replying said that the projects have generated unanticipated spin-offs, and 84 per cent said that the benefits of the products outweigh the costs.

The survey confirmed widespread dissatisfaction with various aspects of EC funding, in particular the time and cost involved in making grant applications to Brussels. "It was clear that the aggregate cost of all those making applications was a significant drag on the benefits to those who were eventually successful", says Georgiou.

It also revealed what many had suspected: that EC funding, like most other areas of research funding, is concentrated in

the more affluent areas of the country. Apart from pockets of activity around cities such as Manchester and Edinburgh, by far the largest proportion of grants goes to south-east England.

One unexpected finding is that only 14 per cent of grant recipients describe their project as strategic research, a description that corresponds to what Brussels, nearly a decade ago, began to call "precompetitive" research. The term connotes research on which industrial competitors are supposed to collaborate to produce a common body of knowledge and to protect such joint ventures from the antimonopoly rules of the Treaty of Rome.

In contrast, more than half of the Brussels-funded projects fit the description of applied research or development. "Our overall impression is that there is probably more of a downstream focus than one would have imagined", says Georgiou.

Another finding casting doubt on the idea of precompetitive research is that only 25 per cent of industrial groups involved in such research were collaborating with direct competitors. In contrast, 39 per cent were collaborating with "indirect competitors", those producing

similar goods for different markets.

Some 32 per cent of those polled said that they were collaborating with customers and 19 per cent cited suppliers, confirming the impression that the projects have created a degree of vertical integration in the industries covered. But the most popular motive for collaboration was to gain access to complementary expertise. About 30 per cent of participants said that they intended to collaborate with other research groups on commercial applications.

Although the general attitude of industrial participants to their involvement in EC programmes was positive, it was less enthusiastic than that of academics; only 67 per cent said that the benefits outweighed the costs. Even more significantly, the survey found that only 40 per cent of the 100 companies in Britain that spend the most on research and development were participating on EC-funded projects.

Robert Jackson, parliamentary undersecretary at the OST, said that he was "disappointed" by that result. He has suggested to officials from the Department of Trade and Industry and the OST that they should investigate the reasons for such a poor record of participation. **David Dickson**

Humboldt faculty holds its breath

Munich. The entire academic staff of Humboldt University in East Berlin is anxiously awaiting word on who will be retained as the university becomes the last to apply criteria used in western Germany.

Western faculty are employed according to the Hochschulrahmengesetz, the general law which determines working conditions and pay at universities. Following reunification in 1989, it was decided that academics in east German universities should be employed under the same conditions. Most universities have completed the process, but Humboldt is taking longer — right up to the 31 December 1992 deadline — because of the magnitude of the overstaffing.

At Humboldt, those retaining their jobs will receive pay raises because the new contracts employ staff at 80 per cent of the present western German wage level rather than the current level of 60 per cent. (In 1989, the level was 15 per cent.) The general law also entitles them to greater research funding and better equipment. However, because the number of academic staff is being drastically reduced and more eastern German students are applying to university, the remaining staff will face the growing, nationwide problem of overcrowding.

This state-financed restructuring of staff

at Humboldt is being carried out by a special commission set up after reunification to study the staffing needs of the university. It is redefining some jobs and abolishing posts that have become obsolete since the change in political systems. Anyone from west or east can apply for the jobs, which have been advertised widely.

A commission headed by the president and chancellor of the university makes the final decision on applicants. It also can point the finger at anyone judged to have abused their connections with the East German secret police — although such investigations are also being conducted by education ministers in each of Germany's 16 states (see *Nature* 359, 762; 1992). This means that incumbent staff will have to compete not only with western German scientists but also with their past.

This sudden — and very expensive — change is part of a scheme to impose uniform standards on the employment of academic staff. Although the process may be harsh on the 1,800 staff at Humboldt, many academics are expected to keep their jobs. The university announced last week that 100 people, 44 from Humboldt and 56 from western Germany, have already been recruited for the 550 positions to be filled. **Matthew Beard**

Reforming science in Japan

SIR — Despite a certain partiality (11 references to Tokyo University in 9 pages), some of the observations made by John Maddox and David Swinbanks in their report on "Science in Japan" (*Nature* **359**, 573–582; 1992) were so ironically succinct as to make us both groan and laugh in affirmation. That the rigidity of the system in Japanese universities leads to inbreeding is illustrated by the fact that in the veterinary faculty of the national university where we work, 13 out of 16 professors and all 16 of the associate professors have received at least one degree at this university. No women are employed as staff members, no true postdoctoral positions are available, and one third of this year's graduates will remain for another four years to do their PhD. The number of laboratory technicians has dropped in the past 20 years (currently only one is employed), and grant conditions usually make it impossible to hire them. There are, however, some encouraging signs: two years ago the faculty made its first-ever appointment of a non-Japanese (albeit a graduate) to a staff position.

Twenty years ago in the *Nature* special issue on Japan, attention was drawn to criticisms of problems as well as to overbearing administration and defects of the Monbusho grant system. These problems are all alive and well in Japanese universities today. But one wonders if these conditions are really amenable to reform by policy or decree, or whether, like many of Japan's characteristics, they are a product of compromise between the desire for stability and the necessity for improvement. Any type of reform would be more likely to involve consensus and consultation at all levels, a slow process but one that usually receives support from the majority of the people affected. It must be remembered that Japan has made significant contributions to science despite its shortcomings. The attention to detail, care and diligence of Japanese researchers will ensure that these people will continue to produce top class research, and hopefully to be given the recognition their efforts deserve.

Matthew C. Playford

Masao Kamiya

Department of Veterinary Parasitology,
Hokkaido University,
Sapporo, Japan

SIR — We appreciate the attention that *Nature* pays to Tokyo University, but in the News article (**390**, 403; 1992) about the external evaluation of one of our departments there are two points on which we believe our motives have been misinterpreted.

The first concerns the white paper or policy document (*hakusho* in Japanese), a 630-page report published by the university. Because of its length, the report was embargoed until 7 December to give the press time to digest it. For the same reason, we organized a general lecture on 27 November followed by a press conference on 4 December. Regrettably, *Nature* published comments on the contents of the *hakusho* on 3 December, before the embargo was lifted.

The second point concerns the content. You quote one of the deans of Tokyo University as saying that the *hakusho* is largely "propaganda" and contains little "true evaluation". However, the dean himself has explained that his argument was rather that the *hakusho* is a report of our internal study, which will be followed by an independent external evaluation. In other words, the *hakusho* is the first necessary step to "true evaluation". The complex nature of the task of this internal evaluation demands that we inform people of the process through the contents of the *hakusho*.

M. Aoyagi

(Chairman)

Committee of Public Relations,
University of Tokyo,
7-3-1 Hongo, Bunkyo-ku,
Tokyo 113, Japan

□ *Nature* did not take part in the meetings of 27 November and 4 December and received no material from them. We learned of the *hakusho* independently from several sources. — Editor, *Nature*.

Irradiated graphite

SIR — In 1949, I was sent from the Atomic Energy Research Establishment, Harwell, to the Chalk River Laboratory, Canada, to study the possible changes in the physical dimensions of graphite under irradiation in the NRX nuclear reactor. The main objective was to confirm the existence of a dimensional effect and establish a quantitative relationship between exposure to radiation and dimensional change¹.

After accumulating some graphite samples with considerable exposures, my colleagues and I did some heating experiments to see whether the effect could be annealed out. Clearly, some recovery took place at 100, 200 and particularly 250 °C. In one case the oven temperature appeared to rise much higher than it should have done, possibly higher than the oven design limit; the sample had annealed much more than the others. At this stage we assumed that

the oven thermostat or thermocouple had failed. We were not able to continue these tests because only a few highly irradiated samples were available.

We believed that the effect we were measuring resulted from increasing disorder in the crystal structure caused by carbon atoms being knocked out of their lattice positions by fast neutrons into interstitial positions between the hexagon layers characteristic of graphite. Presumably, the displaced atoms are in positions of greater potential energy. We used the term 'Wigner effect' to describe the dimensional changes. I was not then aware of the significance of stored energy in graphite (later referred to as the Wigner effect). In 1953, it was appreciated that graphite-moderated reactors could store energy in the graphite, following an accidental release in one of the Windscale (now Sellafield) reactors. Following this, the stored energy was released in a controlled way as part of the routine operating procedures.

In October 1957, a controlled release in one of the Windscale reactors went seriously wrong, resulting in a fire in the core, which was destroyed. Subsequently, the graphite research programme was considerably extended.

It seems possible that in our anomalous annealing measurement in 1949–51 we were observing the release of stored energy in the graphite sample. Had we followed our preliminary measurement with more studies, we might have understood the properties of energy storage and release several years before the Windscale accident. Such a view, of course, has the benefit of hindsight.

N. J. Pattenden

73B Essex Street,
Newbury, Berkshire RG14 6RA, UK

1. Sheard, H. & Pattenden, N. J. *The Effect of Pile Irradiation on the Linear Dimensions of AGXP Graphite*, Chalk River report CRNE-496, 1952.

CEA research

SIR — Declan Butler's statement that "research in plant physiology will be phased out during the next three years" at the Commissariat à l'Energie Atomique (CEA, *Nature* **359**, 568; 1992) has caused some concern to researchers here.

In fact, only microalgae biotechnology and radioagronomy are to be phased out in the coming year; research on plant ionization will concentrate on fundamentals and will be expertly appraised; plant physiology and research on artificial ecosystems will remain a priority, as will protein engineering.

Michel Susillon

Commissariat à l'Energie Atomique,
31–33 Rue de la Fédération,
75752 Paris Cedex 15, France

Rational disposal of chemical weapons

Stephen Black and Benoit Morel

The world's chemical weapons are due to be destroyed within 10 years. How sensible is a radical proposal to destroy Russia's stockpiles by means of underground nuclear explosions?

AFTER many years of negotiation, a ban on production, possession and use of chemical weapons was agreed this summer in Geneva. An important provision of the treaty is that countries possessing chemical warfare agents and with chemical weapons deployed on their territories have 10 years to dispose of them. In contrast to the frenzy of other demilitarization efforts since the end of the cold war, destruction of chemical weapons presents significant unresolved technical and economic problems. Solutions appropriate to particular countries may not be universally acceptable because of technical, economic and social differences. In particular, the use of underground nuclear detonations to dispose of chemical weapons has been dismissed in the West but is being actively promoted by a Russian company.

The destruction of chemical weapons is more difficult and more expensive than their manufacture. In addition to their extreme toxicity, the high explosive bursting charge contained in many of the munitions adds a significant degree of complexity. There have been many suggestions for their industrial-scale destruction. The cost, difficulty and efficiency of these approaches vary, but none seems ideal.

Destruction methods

Neutralization of chemical warfare agents is much easier than their complete destruction. The first industrial-scale process used by the US Army was the chemical neutralization of Sarin at the Rocky Mountain Arsenal. The process used sodium hydroxide to produce a significantly less toxic waste product. This method was discontinued for various reasons¹: the sometimes slow reaction rate; the reformation of Sarin in the waste stream; the toxicity and quantity of solid waste; and the cost and complexity of plant construction. Neutralization is not widely accepted as a satisfactory means of disposal because it does not reverse the most difficult manufacturing step, the creation of a phosphorus-methyl bond. This bond has a very high energy of activation, and as a result is very robust². Many specialists believe that it would be too easy to remanufacture nerve agent from the waste product of chemical treatment.

Incineration, the only other industrial-scale process tested, was the method

selected by the US government for the destruction of its arsenal. Despite optimistic forecasts, the projected cost of this programme has risen to around \$8 billion, the disassembly of the weapons continues to cause difficulties, and the programme has adverse effects on the environment. Incineration, as applied at the US Army's facility on Johnston Atoll, should not be considered a universal panacea for stockpile disposal.

Alternative techniques such as controlled explosion of chemical warheads in a pool of sodium hydroxide, cryofracture (cooling the weapon in liquid nitrogen before crushing it in a hydraulic press), supercritical water oxidation, and steam gasification are being studied. Yet none of them has been applied on an industrial scale and each presents technical problems similar to neutralization or incineration.

Another alternative, the use of underground nuclear detonations, presents a unique set of possible benefits but also problems. Peaceful nuclear explosions are being advocated as the most cost-effective or 'rational' approach to the destruction of large quantities of transportable chemical warfare agents, and chemical munitions. The International Chetec Corporation, based in Moscow, is offering to destroy weapons and other hazardous materials in controlled underground explosions³. The two most obvious advantages of this approach are first, that given the physical conditions created by a nuclear explosion large quantities of chemical agent can be completely atomized in an instant, and second, there would be no need to dismantle live munitions, which is the most delicate and expensive phase of chemical warhead destruction. But the proposal raises significant questions about munition transport, emplacement and environmental impact.

For primarily political reasons the idea seems to be a nonstarter in the West. But it is a option for other countries, particularly Russia, which has inherited around 50,000 tons of chemical warfare munitions. The existing Russian programme for the destruction of its chemical arsenal is meeting various obstacles. Given the cost of building safe incineration facilities, the dire conditions of the Russian economy and the low initial investment, the nuclear option may be attractive in this case.

Underground explosions whose purpose is to destroy chemical weapons differ considerably from weapon tests. The size of the hole in which the explosion has to take place is somewhat larger in the latter. This not only adds to the cost, but also affects the dynamics of the explosion and its coupling to the surrounding rock. Also, the atomization of the chemical agent leads to the production of large quantities of noncondensable gas.

Explosion effects

The dynamics of a 'tamped' underground detonation are reasonably well understood, and provide a good starting point for assessing the use of nuclear explosions for destroying chemical weapons^{4,5}. At detonation, roughly 70 tons of material per kiloton of yield is vaporized⁶, resulting in the creation of a cavity containing compressed gas at very high temperature. The energy communicated to the rock surrounding the cavity leads first to its fracture and, as the shock wave abates, to the generation of an elastic deformation propagating as a seismic wave. Once the cavity has reached its maximum diameter it recompresses under the effect of an elastic rebound. The cavity diameter shrinks until the hoop stress generated by radical convergence approaches the compressive strength of the rock. At the end of this dynamic phase, which lasts from 50 to 80 milliseconds, the cavity has reached a metastable equilibrium between various forces: the high-pressure gas inside, lithostatic pressure from the overburden, and compressive hoop stress.

The effect of the shock wave depends on the geology of the surrounding rock. In the porous rock of Nevada, the rebound creates a 'containment cage' of fractured rock surrounding the cavity in which compressive hoop stress is at its maximum. The containment cage thickness is typically 60 metres per $\text{kT}^{1/3}$. But in granite, where the tests in the former Soviet Union were conducted, the rock does not fracture in the same way and the containment cage effect is less pronounced.

The containment cage usually prevents the release of radioactive contaminants by reducing the two prime channels of containment seepage: diffusion through the surrounding rock and flow through natural or explosion-induced fissures.

This is one explanation of why the Soviet tests tended to generate more seepage than their US counterparts. The methodology used in the United States to ensure containment and avoid seepage has been quite effective. It is based on the rule of thumb that burial at 122 metres per $\text{kt}^{1/3}$ will prevent surface contamination.

The mass of chemical munitions to be destroyed must be scaled to the yield of the explosion in such a way that there is no possibility of any agent surviving the initial energy release. To make sure that all the agent is thoroughly destroyed we can use the rule of thumb that a 100-kiloton detonation will vapourize roughly 7,000 tons of surrounding material. With such a yield it is 'safe' (maybe a bit conservative) to assume that 5,000 tons of chemical shells (1,000 tons of agent and 4,000 tons of shell steel) can be destroyed.

At the end of the dynamic phase, after the rebound creating the containment cage, 7,000 tons of gas will be trapped in the cavity where, according to another rule of thumb, 35% of the energy remains⁶. The vapourization of 4,000 tons of iron requires roughly 8 kilotons of energy. If approximately 20 kilotons is used to heat up the gas trapped in this cavity, it should reach a temperature of around 10,000 kelvin and a pressure between 40 and 100 atmospheres. The pressure and temperature at the end of the dynamic phase should not significantly differ between a tamped weapons-test explosion and a disposal detonation. However, the gas composition of the cavity will be markedly different in the two cases, as we will show in a more detailed study (manuscript in preparation). The main difference is the creation at the end of the cooling phase of far more light non-condensibles like hydrogen or methane when chemical weapon munitions are destroyed by the explosion.

In both types of detonations the amount of hydrogen generated is basically the same. In a tamped shot, far more oxygen is created which can recombine during cooling with the hydrogen to make steam or water. However, the atomization of chemical shells leads to the generation of large quanti-

ties of iron and carbon which will avidly wash away the oxygen leaving the hydrogen either to become gaseous hydrogen or to recombine with the remaining carbon to form light noncondensable hydrocarbons like methane. The appearance of substantial quantities of light noncondensibles in the cavity changes the conditions for transport of gas and radioactive contaminants after the explosion.

Under a pressure gradient, the different components of a gas mixture travel in porous media, at different velocities⁷. The flux of the light components is inversely proportional to the square root of the molecular mass of the constituent. The consequence of the resulting transport at high velocity of light constituent is, all other conditions being equal, an increase of 20 to 40% in the flux of the heavier radioactive products either gaseous (krypton⁸⁵, tritium and tritinated methane) or particulate. One way to avoid radioactive surface contamination would be to increase the depth at which the explosion takes place by a similar proportion, but this would be expensive. There is also a constraint placed on burial depth by the location of the water table.

Of further concern is the fact that underground nuclear explosions for chemical weapons destruction could produce more hydrofracturing than tamped explosions. Hydrofractures are cracks through which contaminants can migrate and possibly reach the surface. Gas and particulate transport through these fractures, when they exist, is the most dominant contaminant transport mechanism for underground gas flows⁸. The size of a given crack is dependent on a balance between compressive stress in the fractured material and the gas pressure propagating the crack⁹. A decrease in compressive stress or an increase in cavity pressure will increase the size and extent of hydrofractures.

The explosion of a nuclear device in a cavity filled with chemical weapons could correspond to a 'partially decoupled' explosion¹⁰. An explosion is said to be decoupled when a significant part of the released energy is not communicated to the rock. As a result there is not enough residual hoop stress to create a

good containment cage¹¹. The nuclear destruction of chemical weapons could lead to a partially decoupled explosion. The cavity must be large enough to accommodate the chemical munitions and a portion of the explosive energy is used to vapourize the weapons and thus will not contribute as much to the creation of the containment cage. There is therefore a serious possibility that weapon destruction shots could result in a less efficient containment while at the same time producing more noncondensable gas.

Policy implications

Is the destruction of chemical weapons through underground nuclear explosions unsafe and unadvisable? Such a conclusion does not necessarily follow from our findings, but it is certainly true that the presence of chemical weapons changes the physics and the chemistry of the explosion and its containment. We have shown here that the rules of thumb currently used to contain underground nuclear explosions do not achieve the same level of safety if applied to explosions in the presence of chemical weapons.

Past experience with gaseous radionuclide seepage suggests that the danger represented by some marginal seepage may not be great. Even though it occurs to a much lesser extent, the effects of particulate transport can be much more dangerous. It would thus be imprudent to declare nuclear destruction of chemical weapons a safe approach. The cost efficiency is also questionable, as each explosion does not seem to destroy a large amount of chemical munitions. The technique requires a large hole deep underground and the associated risks of munition emplacement. It supposes also the ability to bring the munitions to the test site.

Yet nuclear explosions must be compared with the alternative, which can also be expensive, complicated, preliminary and polluting. The conclusion may be that the nuclear destruction of chemical weapons is probably not as attractive as it may have looked superficially, nor as clean or cost effective. Yet it may still be the most appropriate course for a country like Russia which has significant economic constraints, a very large chemical arsenal, and which is bound by treaty to destroy its weapons within a certain time. □

Stephen Black is at the Graduate School of Public and International Affairs, University of Pittsburgh, 4623 Forbes Quadrangle, Pittsburgh, Pennsylvania 15260; Benoit Morel is at the Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890, USA.

1. US Department of the Army Chemical Agent and Munition Disposal: Summary of the U.S. Army Experience (Office of the Program Manager for Chemical Demilitarization, Maryland, 1987).
2. Black, S., Morel, B. & Zapf, P. *Nature* **351**, 515-516 (1991).
3. Andryushin, I. A., Trutnov, Yu. A. & Chernyshev, A. K. Report on some Aspects of Creation of Nuclear Explosion Technology for Destruction of Toxic and Hazardous Materials and Waste Products (Arzamas 16, 1992).
4. US Office of Technology Assessment The Containment of Underground Nuclear Explosions (Washington, DC, 1988).
5. Johnson, G. W., Higgins, G. H. & Violet, C. E. J. *geophys. Res.* **64**, 1457 (1959).
6. Duff, R. *The Feasibility of Chemical Munitions Disposal*

Using Nuclear Explosions DNA-TR-81-205 (Defense Nuclear Agency, 1982).

7. Mason, E. A. & Malinauskas, A. P. *Gas Transport in Porous Media: The Dusty Gas Model* (Elsevier, Amsterdam, 1983).
8. Burkhard, N. R., Hearst, J. R., Peterson, E. W., Nilson, R. H. & Lie, K. H. *Containment of Cavity Gas in Fractured or Rubblized Emplacement Media* UCRL-100953, DE90001054 (Lawrence Livermore National Laboratory, 1989).
9. Kunkle, T. D. & Travis, B. J. *Hydrofracture From a Growing Cavity* LA-UR-82-2308 (Los Alamos National Laboratory, 1981).
10. Patterson, D. W. J. *geophys. Res.* **71**, 3427 (1966).
11. Glenn, L. A. *Energy Density Effects on Seismic Decoupling* UCRL-JC-109393 (Lawrence Livermore National Laboratory, 1992).

Earth saved from disaster!

There were high hopes that comet Swift-Tuttle might be on collision course with the Earth, but it now seems that enthusiasts for destruction will have to wait for another ill-omened body.

THE profound satisfaction that comes from contemplating the destruction of our entire planet is presumably an infinitely multiplied form of the satisfaction we get from watching the bad guy's car, at the end of a movie, speed over the edge of a cliff, plunge into a rocky chasm, and burst into flames. We are alarmed, for example, at the thought that human activity might be changing the climate of the Earth enough to cause the demise of our species, but at the same time is there not a certain thrill in the idea that we might have become so technologically capable that we have the power to extinguish a whole planet?

For the connoisseur of catastrophe, however, global warming is a protracted and uncertain venture. Already there are moves afoot to curtail our collective output of greenhouse gases into the atmosphere, and it seems quite possible that global warming may be averted before anything genuinely destructive comes of it. Global destruction, if it is to have wholesale appeal, should be abrupt, and should have a degree of cinematic splendour. This is why, surely, there is so much continuing fascination with the extinction at the Cretaceous/Tertiary boundary and the end of the dinosaurs. A slow dwindling of the dinosaur population, perhaps because of changing climate, is unappealingly tame, but if an asteroid or comet ploughed into the Earth and sent up dust that darkened the skies and snuffed out a substantial fraction of all life on the globe, then we have a screenplay our imaginations can latch onto. Of course, we are still here to talk about the dinosaurs, and there is an extra vicarious pleasure to be gained from thinking that we — in the form of some small and adaptable ancestral mammal — were smart enough to survive the destruction that did for the dim-witted and lumbering reptiles.

In science-fiction books and movies that portray an Earth in the grip of some imminent death-threat, there is generally at least one smart and adaptable creature who, mocked at first (just as the first rodents were no doubt scorned endlessly by the dinosaurs), comes through with a brilliant scheme to save the world — excluding those who mocked, of course. Earlier this year, it was suggested in the United States that a good way to adapt the defence industry to modern times would be to have the ingenious weapons designers work on saving the Earth; their expertise could be used for a massive effort against global warming or, more excitingly, they could be asked to find ways to

make nuclear weapons effective against external threats, notably the impacts of comets or asteroids. These suggestions were largely discounted as either special pleading for the defence industry or foolish romancing by scientists who had not yet noted that the world's economy was in recession. But of course it has always been the fate of the farsighted to be shunned, and anyway astronomers could prove that a collision between the Earth and a chunk of celestial rock large enough to be damaging must, by the laws of probability, happen one day.

In happy fulfilment of these dire thoughts, comet Swift-Tuttle entered onto the scene in September. The purely scientific interest in this matter was that Swift-Tuttle, first observed in 1862, had been predicted to return in 1981, but never showed up. But Brian Marsden, of the Harvard-Smithsonian Center for Astrophysics, had also suggested the possibility of a later return, on the grounds that a comet seen in 1737 might have been Swift-Tuttle on an earlier circuit. The 1981 return was predicated on an orbit governed purely by gravitation, but to tie the 1737 and 1862 apparitions together, Marsden had to throw into his calculations some non-gravitational effects: comets eject gas when they are heated by the Sun, and these jets can be enough to perturb the orbit. With an estimate of non-gravitational effects included, Swift-Tuttle's orbit would, according to Marsden, bring it back to the inner solar system in 1992. In September an amateur comet-watcher in Japan spotted an object in Ursa Major which the experts soon declared to be Swift-Tuttle.

So much for the science. What made Swift-Tuttle exciting was that its orbit, recalculated by Marsden with the help of the new sighting, will bring it very close to the Earth on its next return, in 2126. In fact it will miss the Earth by about 19 days, or 15 million miles, which is close enough to be interesting but hardly the stuff of which movies are made. Why then did this modest prediction turn into newspaper stories foretelling the possible terrestrial impact, a century and a half hence, of a celestial body comparable to the one that wiped out the dinosaurs? What prompted the editors of *Newsweek* to run a cover story entitled "Doomsday Science", complete with artist's rendition of Swift-Tuttle bearing fiercely down on an innocent, pearly-blue Earth?

When he reported his calculation in the IAU Circulars (a telegram system by which astronomers quickly alert each other to important observations and events), Marsden

noted that the chance of a collision of Swift-Tuttle with the Earth, in 2126, was about one in ten thousand. But he also observed that the non-gravitational influences on the orbit, sufficient to shift the predicted return by 11 years, from 1981 to 1992, were inherently unpredictable, and that a mere 15 day correction to the calculated perihelion of Swift-Tuttle would be enough to put it on a collision course with Earth.

It is tempting to think that if the time of perihelion can change by 11 years, then a change of merely 15 days is somehow a likely possibility, but of course the opposite is true: if the moment of perihelion is random in the interval plus or minus 11 years around the predicted return, then the chance of Swift-Tuttle arriving on any one chosen date is indeed about 1 in 10,000. The *Newsweek* article noted the low probability of collision but went on to list in enthusiastic detail the wondrous horrors that would ensue: an explosive force of 100 million megatons of TNT, a plume of vaporized stone, buildings and even trees bursting into flame, rain as caustic as the acid in a car battery, and other phenomena marvellous to imagine.

Unfortunately for catastrophists, none of this seems at all likely, and there is little chance that reformed weapons designers will be able to show off their talents by designing missiles that would intercept Swift-Tuttle and push it off track by means of artificially induced non-gravitational forces. IAU Circular number 5672 however offered a double tease, briefly raising hopes of a fatal collision. First there was a suggestion that Swift-Tuttle, because it suffers unusually large non-gravitational perturbations, might also be susceptible to fragmentation; if a shower of large comet chunks filled the orbital path, Earth's chances of intercepting a dangerously large rock would be correspondingly larger. Second on Circular 5672 was an observational note that Swift-Tuttle indeed seemed to have split into pieces — three secondary nuclei were spotted, travelling *en famille* with the parent body. Gratifyingly, the odds of doomsday were on the up again.

Alas, IAU Circular 5673 squelched even these hopes. It reported that the observations of secondary nuclei were due to optical reflections in the telescope, and that Swift-Tuttle was therefore still a single comet plodding to its non-*rendezvous* with the Earth. Despite all the hopes pinned on it, Swift-Tuttle is resolutely refusing to be the agent of wild destruction. Too bad!

David Lindley

Dark matter versus magnetism

James Binney

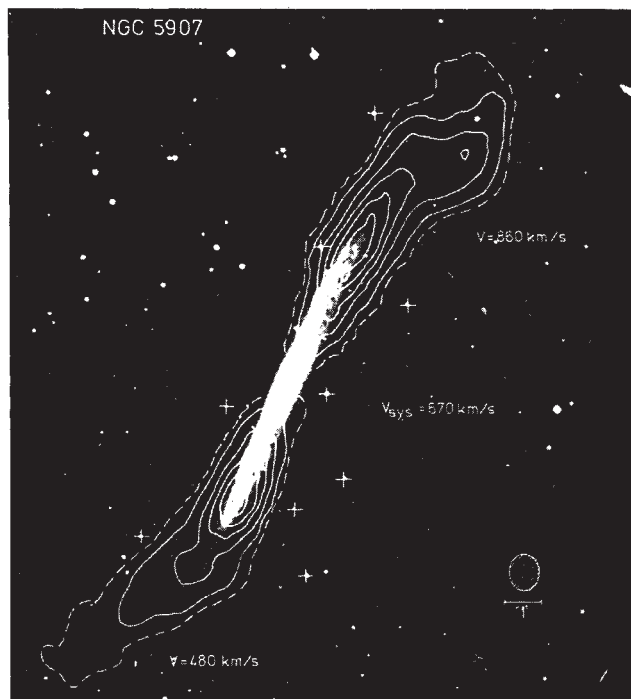
THE strongest evidence for the existence of dark matter in the Universe comes from the rotation velocity of gas orbiting in spiral galaxies: this does not fall off with increasing galactocentric radius as quickly as one would expect from the visible distribution of mass. Elsewhere in this issue (*Nature* **360**, 652–653; 1992), E. Battaner *et al.* suggest that this evidence may be flawed because it fails to take into account the effects of magnetic stresses within the observed gas.

Interstellar gas generally contains sufficient free electrons to act as a near perfect conductor, able to sustain the currents to support a magnetic field. Also, from the alignment of dust grains, a few measurements of Zeeman splittings of spectral lines, the observation of synchrotron emission by relativistic electrons and from the observation of Faraday rotation of polarized radio emission, we know that interstellar gas is permeated by a magnetic field. A field of 1 picotesla (10 microgauss) or more would exert forces on the gas sufficient to affect its dynamics significantly.

Back in the 1950s it was widely believed that spiral structure in galaxies was a predominantly magnetic phenomenon, as were the striking intergalactic filaments seen near strange galaxies such as NGC4038/9. But from the mid-1960s onwards it became increasingly clear that gravity acting alone could account for spiral structure; and intergalactic filaments turned out to arise naturally when a disk galaxy is tidally distorted by a close companion (J. Barnes and L. Hernquist, *A. Rev. Astr. Astrophys.* **30**, 705–742; 1992). So the conventional position nowadays is that although the interstellar magnetic field probably plays a key role on parsec scales and less, on galactic scales it amounts to no more than a small perturbation on fundamentally gravitational dynamics.

Battaner *et al.* challenge this consensus by arguing that tens of kiloparsecs from the centre of a galaxy, where the gas density is usually very low, the dynamics of the gas may be strongly affected by a magnetic field of the order of 1 picotesla. We know that the interstellar field tends to wind around galactic disks, so that to a first approximation one can imagine that the field lines form

a series of gas-filled hoops within the plane of the disk. The magnetic field B imposes two kinds of forces on the hoop: a tension $\frac{1}{2}B^2/\mu_0$ runs the long way around the hoop, pulling it more snugly around the galaxy, while in the two perpendicular directions, the field exerts pressure $\frac{1}{2}B^2/\mu_0$, tending to cause the hoop to fatten into a doughnut (where μ_0 is the permittivity of free space).



Optical image of the stellar disk of the galaxy NGC5907, with radio-emission contours showing hydrogen flaring at the edges — a sign of magnetic tension?

The magnetic tension in the hoop must be counteracted by some combination of a net outward force on the hoop due to a radial gradient in the perpendicular magnetic pressure, and an excess of the centrifugal force on the spinning hoop over the inward gravitational pull of the galaxy. Battaner *et al.* deduce the field configuration within the Andromeda galaxy, M31, that would cause the gas to rotate at the observed speeds at each radius if the only matter in the galaxy were that associated with the observed stars. They find that a field of magnitude 0.8 picotesla from 10 to 30 kiloparsecs does the trick.

How might one confirm that a field of this magnitude is present? From the alignment of dust grains one learns the orientation but not the magnitude of the field, and the field is much too weak to be detected from a Zeeman splitting. This leaves Faraday rotation and synchrotron emission as possible probes.

Synchrotron emission is proportional to the product of the energy densities $B^2/2\mu_0$ and ϵ in the field and in relativistic electrons. Battaner *et al.* obtain a reasonable fit to the synchrotron radiation observed at large radii from M31 by assuming that ϵ scales with the gas density ρ as ρ/B ; the motivation for this assumption is that the source of the relativistic electrons is the gas, and synchrotron losses due to B constitute the sink. But this calculation does not form a rigorous test of the theory in that the normalization of the density of relativistic electrons is set by requiring that

theory and observation agree at 10 kiloparsecs.

In our own Galaxy, Faraday rotation of radiation from pulsars has been used to probe the magnetic field along many lines of sight. The conclusion of these studies (L. Spitzer *Physical Processes in the Interstellar Medium*; Wiley, New York, 1978) is that near the Sun the azimuthal field is around 0.2 picotesla, a factor of four smaller than the field required by Battaner *et al.* in the outer regions of M31. As magnetic forces are proportional to B^2 , these forces are locally more than an order of magnitude smaller than those required by Battaner *et al.*

Unfortunately, in regions significantly further from the Galactic Centre than we are, the azimuthal field makes only a small contribution to measurements of Faraday rotation, because it runs almost perpendicular to the relevant lines of sight. Such measurements, therefore, can probably

not rule out the possibility of an outwardly increasing azimuthal field of the type required by Battaner *et al.* But studies of the vertical structure of the interstellar medium should allow one to place a firm upper limit on the dynamical significance of the azimuthal field. Indeed, if the azimuthal tension associated with B is strong enough to make a serious contribution to the centripetal force, the perpendicular pressure will prevent the gas disk being strongly flattened. Hence, as one moves out into the region of dynamically significant field, the gas disk should flare very strongly. Observations of other galaxies show that galactic disks *do* flare at the edge of the galaxies' optical images (see figure). But do they flare as rapidly as the model of Battaner *et al.* requires? □

James Binney is in the Department of Physics, Oxford University, 1 Keble Road, Oxford OX1 3NP, UK.

Genes map the migratory route

William J. Sutherland

ON page 658 of this issue¹, Berthold *et al.* show that the blackcap *Sylvia atricapilla* has undergone dramatic changes in migration route over the past 30 years and, moreover, that these changes have a genetic basis. The migration routes taken by birds must have altered in response to variations in climate and the environment such as those that prevailed during the ice ages. Why, though, has this quick and recent alteration of route occurred?

Blackcaps used to be seen very rarely in Britain in winter, but were common during the summer breeding season. Since the 1960s, however, the wintering population has markedly increased, and these birds are now regularly seen, particularly in suburban gardens². The overwintering birds spend the early winter feeding on natural food such as berries, but once these are depleted they live off the wide range of food provided on bird tables and can maintain or even increase their weight during the winter³.

At first it was assumed that some of the British breeding blackcaps had simply decided to take up permanent residence. However, studies of birds ringed in Britain and mainland Europe during the breeding season indicate that these overwintering birds are not British blackcaps extending their stay, but birds from Germany and Austria that have pioneered a new north-westerly migration route which results in their wintering 1,000–1,500 km north of their traditional wintering area around the western Mediterranean. None of these northwest migrants had been recorded in Germany or Austria before 1961, despite the thousands of birds that were ringed there; but in parts of these countries they now account for 7–11 per cent of the breeding population.

To study this switch in orientation, Berthold *et al.* captured some birds wintering in western England and returned them to Germany. The standard way of recording migratory orientation is to record the activity of the birds in circular cages at night during the migratory period, a response known as migratory restlessness (the direction and duration of migratory restlessness correlates

well with actual migratory patterns). Not surprisingly, the British birds showed migratory restlessness in a north-westerly direction in the autumn, whereas those from a population in southwest Germany showed a southwest preference, in accordance with their migration to the western Mediterranean.

To distinguish whether the basis of this unusual behaviour is genetic or ac-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Winter holiday — why has the blackcap changed its migratory direction?
(Picture by K. Wothe.)

quired, Berthold *et al.* allowed the British birds to breed with each other and then observed the behaviour of the offspring. The migratory restlessness of the offspring was also directed to the northwest, showing that the switch in orientation has a genetic basis. Furthermore, the orientation behaviour of siblings was more similar than that of unrelated young, which suggests that there is a genetic component to the variance in orientation.

Earlier work by Berthold and his colleagues⁴ has shown that both the direction and extent of migratory behaviour has a strong genetic component, and breeding experiments indicated that rapid changes in behaviour might be possible. Their present study shows that dramatic changes in migration can actually be accomplished within decades.

For such a rapid change to take place, the selection pressure must have been strong. One reason for the switch to the

northwest route could be to take advantage of the increasing habit of providing food on bird tables. Alternatively, winters in Britain have become milder over the past 30 years⁵, so it is possible that the switch in migration pattern is a response to this.

Blackcaps exposed to photoperiods simulating those in Britain started their migratory activity earlier than those exposed to the photoperiods that would be expected for birds using the traditional wintering grounds in central Spain⁶. This suggests that birds wintering in Britain may return to the breeding grounds about ten days earlier. If so, this is likely to provide an advantage in gaining territories and may have led to assortative mating (like breeding with like), which might in turn have contributed to the rapid spread of this new migratory behaviour.

It is likely that bird migration patterns have been changing for thousands of years as a consequence of man's activity. Another example⁷ is shown by the behaviour of the sand martin *Riparia riparia*, which winters in Africa. After the introduction of the predatory Nile perch *Lates niloticus* to Lake Victoria in 1959–60, the native cichlid fish population crashed; this resulted in a population explosion of emergent insects, which at times resembled a cloud over the lake. The population of sand martins using the site subsequently increased massively from small numbers to hundreds of thousands.

The need for such changes in migration behaviour is likely to increase as a result of global warming. Berthold⁸ speculates on the likely consequences for European birds. The species in Europe that are likely to benefit are the resident species whose populations are depressed by severe winters. It is also likely that populations of partial migrants will be composed of a greater proportion of sedentary individuals and that short-

1. Berthold, P., Helbig, A. J., Mohr, G. & Querner, U. *Nature* **360**, 668–670 (1992).
2. Lack, P. *The Atlas of Wintering Birds in Britain and Ireland* (Poyser, Carlton, 1986).
3. Langslow, D. R. *Bird Study* **26**, 239–252 (1979).
4. Berthold, P., Mohr, G. & Querner, U. *J. Ornithol.* **131**, 33–45 (1990).
5. Department of the Environment *The UK Environment* (HMSO, UK, 1992).
6. Terrill, S. B. & Berthold, P. *Oecologia* **85**, 266–270 (1990).
7. Møller, A. P. *Svalerne* (Naturhistorisk Museum, Århus, Denmark, 1992).
8. Berthold, P. *Acta Congr. internat. Ornithol.* **10**, 780–786 (1991).

and medium-distance migrants will be selected to migrate less far. After the recent mild winters in Europe there have been many reports from central Europe of short-distance migrants wintering in their breeding areas. The success of the late-arriving, long-distance migrants depends on the degree of competition from resident birds and from short-distance migrants that have already established territories. Increases in the number of

resident birds are likely to be deleterious for long-distance migrants. With further man-induced environmental changes on the cards, the future of many species could well depend on their ability to adapt in the way the blackcap seems to have done. □

William J. Sutherland is at the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

CHEMISTRY

Vocabulary for fuzzy symmetry

Patrick W. Fowler

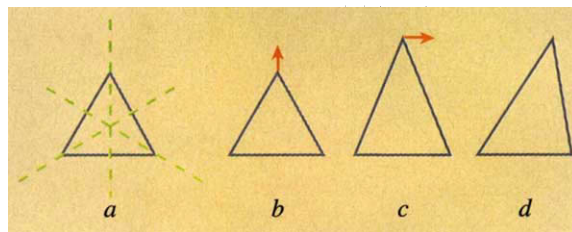
SYMMETRY arguments are standard tools of the trade in theoretical chemistry and physics. Their power lies in the way they deliver reassuringly definite predictions. Sulphur hexafluoride (SF_6) is centrosymmetric and octahedral at equilibrium. Can it support a permanent dipole moment? (No.) Are all six fluorine nuclei magnetically equivalent? (Yes.) Is the molecule optically active? (No.) Should its breathing vibration give rise to a peak in the Raman spectrum? (Yes.) Group theory^{1,2} has all the answers. However, the world is rarely so simple, and Zabrodsky *et al.* have developed a system of "continuous symmetry measures"³ to deal with the murky area of nearly symmetrical or slightly asymmetrical objects. They argue that we should add words like perhaps, maybe and sometimes to the yes-or-no, now-or-never vocabulary of point-group symmetry.

The new approach can be illustrated with the simplest plane figure, the triangle. A rigid, equilateral triangle lying in the horizontal plane has various elements of symmetry. It can be brought into a physically indistinguishable configuration (one in which no interparticle distance is altered) by rotations of 120° clockwise or anticlockwise about a vertical axis, 180° about one of three horizontal axes, reflection in the horizontal plane, or reflection in one of three vertical planes. Isosceles triangles have only two reflection planes and one two-fold axis; scalene triangles have only the horizontal reflection plane. In the notation used by chemists, the three objects belong to the three distinct point groups D_{3h} , C_{2v} and C_s , respectively.

The question addressed by Zabrodsky *et al.* can be put this way: what happens to the symmetry description if we start with an equilateral triangle and gradually distort it? Nudge one of the vertices along a radius and the triangle is now isosceles. Nudge the vertex along a tangent and the triangle becomes scalene. For a traditionalist, all three cases are quite different and are described by

distinct groups, because a symmetry element is lost as soon as the vertex starts to move.

In the continuous approach, the three triangles are described by a gradual movement away from symmetry elements, and each instantaneous con-



An equilateral triangle (a) has four axes and four planes of symmetry. Small displacements (b,c,d) have a catastrophic effect on the name we attach to the triangle and its symmetry group, but may hardly affect the properties.

figuration is characterized by numerical measures $S(C_3)$, $S(C_2)$, $S(\sigma)$ of the deviation from a rotational or mirror symmetry. The procedure for assigning these indices is straightforward, depending on a weighted sum of squares of distances between points and their images in a hypothetical symmetrized set representing the nearest object with the desired symmetry. Things get slightly more complicated in three dimensions, because it is not obvious where the missing axes or planes should be, but a preliminary application to three dimensions is given in the paper. A remaining technical issue is whether the continuous symmetry measure (CSM) can be extended to describe deviation from a point group rather than from a single element.

Why are these apparently trivial geometrical manipulations significant for chemistry and molecular physics? Mainly because chemists still use a largely mechanical model of the molecule. The story is that electrons are light and nuclei are heavy. The nuclear framework vibrates, rotates and translates, held together by electronic glue and electronic springs, and where the nuclei go the electrons are sure to follow. At

each nuclear geometry the electron distribution can be found by solving the Schrödinger equation with clamped nuclei. Breakdown of this simple Born-Oppenheimer picture can be important in certain situations, but the approximation does describe most molecules in their ground electronic states.

A vibrating nonlinear molecule is continually sampling geometries of less than the nominal point group symmetry, and one use of CSM indices will be in parametrizing the property surfaces that describe the new electronic properties that arise in the anisotropic non-equilibrium geometries. An example would be the dipole moment of SF_6 induced by stretching one bond. CSM can also represent the small changes in equilibrium geometry caused by chemical substitution, for example, the (slight) loss of sixfold rotational symmetry in the benzene ring of toluene³. This can already be done using internal displacement coordinates based on a single high-symmetry structure, but CSM seems to

offer an unbiased way of representing simultaneously the relationship between a given structure and a whole set of conceivable symmetric forms. It will be interesting to see if its proponents can link the CSM concept to the group theory of floppy molecules², and to the study of chirality measures⁴.

One case that the new scheme does not seem to

cover is where the symmetry of the electron distribution is lower than that of the nuclear framework. For example, when a rigid molecule is polarized by a uniform field the electron cloud can be strongly distorted, but the nuclei may hardly shift. The molecule has certainly lost symmetry in this process, but in a way that is invisible to CSM.

The problem of any symmetry-based approach is that a broken symmetry implies the existence of a property, but cannot predict its value. Continuous symmetry is a partial answer to this problem, in that a small CSM index implies a 'small' value of the property, but with a constant of proportionality that remains to be established by calculation or experiment in individual cases. □

Patrick Fowler is in the Department of Chemistry, University of Exeter, Exeter EX4 4QD, UK.

1. Cotton, F. A. *Chemical Applications of Group Theory* 3rd edn (Wiley, New York, 1990).
2. Bunker, P. R. *Molecular Symmetry and Spectroscopy* (Academic, Orlando, 1979).
3. Zabrodsky, H., Peleg, S. & Avnir, D. *J. Am. chem. Soc.* **114**, 7843-7851 (1992).
4. Buda, A. B., Auf der Deyde, T. & Mislow, K. *Angew. Chemie Int. Ed. Engl.* **31**, 989-1007 (1992).

Horrible plant species

Jared M. Diamond

AN animal species is usually defined as an interbreeding community of populations that is reproductively isolated from other such communities. This is the so-called biological species concept (E. Mayr *Animal Species and Evolution*, Harvard University Press, 1963). Some botanists claim that this definition is unworkable for plants, because of an allegedly high frequency of phenomena rare in animals, such as asexual reproduction, massive hybridization and polyploidy. But such claims have been based on anecdotal horror stories, not on surveys of the actual frequencies of these horrors. Do the horrors really affect such a high proportion of plants as to undermine the above species definition, or are they just an occasional nuisance? This question has at last been tested by a systematic analysis of an entire local flora, at the instigation of the zoologist Ernst Mayr (*Am. J. Bot.* 79, 222–238; 1992).

As background, recall the long-standing uncertainty even among zoologists over the nature of animal species. Species were initially delineated on purely morphological grounds, with the result that Linnaeus separately named male and female mallards and adult and immature goshawks. When Linnaeus realized that these different plumages were just phenotypic variants related to sex or age, he assigned them to a single species and thereby implicitly accepted the modern definition of species based on interbreeding. It was later realized that named individual and geographic variants also fell within the same species definition. Although taxonomists were the scientists who thus defined species, species also proved an indispensable unit of analysis for ecologists, ethologists and other biologists.

Thus species, like brothers, have two attributes that are often misunderstood. First, species do correspond to a biological reality; they are not mere arbitrary subdivisions of a continuum. Second, species (and brothers) are meaningless when considered alone: they are defined with respect to other species. Surveys of local faunas show that the vast majority (more than 99 per cent) of animal species are well delineated by this criterion.

The local flora used for the corresponding botanical test consists of the indigenous vascular plants of Concord Township in Massachusetts in the United States. For the analysis the Concord flora published by R. J. Eaton (*A Flora of Concord from Thoreau's Time to the Present Day*, Museum of Comparative

Zoology, 1974) has been updated by botanists Ray Angelo and Carroll Wood, with input from other botanists (especially G. Ledyard Stebbins) on chromosome numbers and other problems. The result is a flora of 838 described species available for analysis.

The Concord flora turns out to exemplify all the types of horror stories savoured by botanists. The first horror — asexual reproduction (apomixis) — is

three of these phenomena horrify taxonomists because they are difficult to force neatly into any species definition, whether a definition based on morphology or on reproductive isolation.

Set against this maximum of around 57 (= 17 + 37 + 3) aggravations are 616 described Concord species that pose no problems of separation by either species definition. A further 84 'species' pose no problems because they were described solely on the basis of claimed variation in chromosome number, but the claim proved incorrect (72 cases!) or else valid only for locations distant from Concord township (12 cases).

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The blackberry, a member of the genus *Rubus*, which is notorious for taxonomic difficulties posed by asexual reproduction. (From *Illustrated London Almanack*, 1876/Mary Evans.)

evinced by 17 described Concord species each of which consists of a cluster of clones. Some of the clusters are well enough delineated from each other to pose no problems for a purely morphological species definition, and they may also constitute ecological units analogous to sexually reproducing species. But they surely violate a species definition based on sexual interbreeding.

Another horror (up to 37 described Concord species) consists of autopolyploids, that is, taxa defined on the basis of having more than the normal number of complete sets of chromosomes (for example, four rather than two sets). Some of these taxa can be distinguished from the corresponding diploid taxon in morphology, ecology and physiology, whereas others are indistinguishable except in chromosome number. The autopolyploids also vary in viability and in ability to interbreed with the diploid form. Hence some may, and some may not, be functionally equivalent to species. Still a third horror (about three described Concord species) involves swarms of phenotypically variable hybrids between other named taxa. All

The status of all the other named Concord species can be unequivocally settled by applying the species concept based on reproductive isolation, although many of them are problematic for a morphological species concept. Two groups pass the test for valid species with flying colours: seven sibling species and nine allopolyploids. Sibling species are reproductively isolated but morphologically so similar to other named species that specialists may have difficulty distinguishing them. From the point of view of the plants themselves, these are perfectly normal species that have no problems recognizing themselves; the sole difficulty is that botanists are not as well tuned in to their distinctions. Allopolyploids combine the chromosome sets of two other species and happen to have arisen from two rather than one parental species. However, they behave as perfectly normal species in their reproductive and ecological isolation.

Named Concord species that flunk the test include 26 aneuploids, 24 intrapopulation variants, and numerous rare hybrids. So-called aneuploid individuals

have variable chromosome numbers but interbreed with individuals with normal chromosome numbers. They would never even have been named as 'species' if a cytogeneticist had not counted their chromosomes and inappropriately taken the number as the basis for naming a species. The 24 named 'varieties' turn out to belong to the same interbreeding population as some named species. They received names of their own merely because they differ by a single conspicuous character (possibly a single-gene variant) or by multiple characters still connected by intermediates to the main stock. Many other 'varieties' present at Concord had been named earlier but were not included in the analysed Concord flora, because botanists had already recognized their status before Eaton's 1974 publication. Still other named taxa had already been recognized to be occasional interspecific hybrids.

Finally, there remain about 15 named Concord species for which it is still uncertain whether they constitute distinct species or just varieties of other species. Their status will undoubtedly be resolved by further study, and they pose no difficulty in principle for a biological species concept.

Thus, the overall result is that no more than 7 per cent (around 57 out of 838 named species) of the Concord flora fits poorly into the biological species concept. This is a higher percentage than usually observed for local faunas of animals, but hardly enough to undermine the utility of the concept. I am aware of several analyses underway of other local floras, demonstrating only a modest incidence of horror stories.

All this is not to deny that distinctive features separate plant from animal species. Polyploidy and occasional interspecific hybridization undoubtedly occur much more often among plants than among animals. Allopolyploidy as a mechanism of speciation occurs among plants but is virtually absent in animals. The reproductive isolating mechanisms of higher plants obviously differ from those of animals: for example, active behavioural choices in the latter, sterility or differing pollinators or flowering seasons in the former. Nevertheless, despite the occasional horror stories, plant as well as animal species are most profitably defined as interbreeding communities of populations. The most important conclusion from Mayr's analysis is that this definition is not merely workable — it offers the decisive insights into problematic cases. □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024-1751, USA.

And still there is no pulsar

Roger A. Chevalier

FOR a supernova that confirmed most astrophysicists' expectations, SN1987A is strangely becoming something of an embarrassment. Despite indications that a pulsar, a spinning neutron star, was created when the progenitor star exploded and its core collapsed, no pulsar has yet been seen. New observations from the Compton Gamma-Ray Observatory¹ show that some — but not all — of the present luminosity from SN1987A is due to radioactive ^{57}Co . But even then, the deficit is not attributed to power from a pulsar. Rather, physical processes in the supernova remnant nebula have apparently stored up energy from an earlier time, when the radioactive sources were more powerful, only to release it now².

Observations of the total luminosity of SN1987A during the first 1,000 days following the explosion on 23 February 1987, and the direct detection of γ -ray lines of ^{56}Co , showed that about 0.07 solar masses ($M_{\odot}$) of ^{56}Co provided the radiative power of the supernova. Between days 120 and 900, the luminosity dropped exponentially with a time constant of about 114 days, corresponding to the decay of ^{56}Co . The ^{56}Co starts as radioactive ^{56}Ni , which is synthesized during the explosion and has a decay time constant of 8.8 days, and ends as ^{56}Fe . ^{57}Ni is expected to be synthesized along with ^{56}Ni in the explosion; it

decays to ^{57}Co which, in turn, decays to ^{57}Fe with a time constant of 390 days. Because this explosive nucleosynthesis mode should be the dominant source of ^{56}Fe and ^{57}Fe throughout the Galaxy, the solar ratio $^{57}\text{Fe}/^{56}\text{Fe} = 0.024$ by mass gives a first estimate of the $^{57}\text{Ni}/^{56}\text{Ni}$ ratio to be expected in the supernova. Detailed calculations for the conditions in SN1987A show that the $^{57}\text{Ni}/^{56}\text{Ni}$ ratio should be somewhat larger than the ratio expected from solar abundances³.

The ^{57}Ni abundance is of special interest because it gives rise to ^{57}Co , which is expected eventually to dominate the power input to the supernova because of its long time constant. Studies of infrared lines of Fe and Co gave an estimate of the $^{57}\text{Ni}/^{56}\text{Ni}$ ratio as 1–2 times the solar value⁴, but a more direct estimate was expected to come from detection of the 122-keV γ -ray decay line of ^{57}Co . Kurfess *et al.*¹, using the satellite-borne Compton Observatory, have now detected that line and the implied ratio is roughly 1.5 times the solar value. But this is not enough to account for the full luminosity of SN1987A.

At early times, the decay γ -rays deposited their energy as heat within the supernova, which led to the prediction that the initial exponential fall in luminosity at ultraviolet through to infrared wavelengths should level off from the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Astronomers' elation at SN1987A (top) in the Large Magellanic Cloud has turned to consternation as its pulsar has failed to appear. (Below, Tarantula Nebula, also in the LMC.)

^{56}Co decay curve because of the presence of longer-lived ^{57}Co . Such a leveling was observed in 1990. The observations were difficult because most of the luminosity was at infrared wavelengths and an extrapolation of the spectrum beyond wavelengths of 20 micrometres was needed. Suntzeff *et al.*⁵ find that the assumption of ^{57}Co power input fits the data through to day 1,444, but that the $^{57}\text{Ni}/^{56}\text{Ni}$ ratio needs to be 5 times the solar value, not 1.5. Neither does power input from a pulsar with a constant luminosity provide a good fit to the data. Although there are uncertainties in the observational estimates, there appears to be a clear discrepancy between the direct power input from ^{57}Co and the total luminosity.

Clayton *et al.*² consider several reasons for the discrepancy. The most plausible explanation is that the timescales for the indirect processes that lead to infrared, visible and ultraviolet radiation are comparable to the evolution time so that energy could be stored in the supernova from earlier times when the energy input was large. Two possibilities are that the timescale for electrons and ions to recombine radiatively increases to the point where the recombination rate falls less rapidly than the radioactive power, or that there is an increased time delay for suprathermal electrons to strike the dust grains that give the infrared emission.

Although the energy-excess problem is not being taken as evidence for a pulsar, the 10-second pulse of neutrinos from the initial explosion, detected on 23 February 1987, indicates that a neutron star did form in the early phase. The subsequent lack of evidence for a neutron star is beginning to be an embarrassment — if one exists, its radiative power must be less than $10^{37} \text{ erg s}^{-1}$ or else it would be apparent⁵. The well-known Crab pulsar and its nebula, for comparison, put out $2 \times 10^{38} \text{ erg s}^{-1}$; and originally, when it was spinning faster, the luminosity might have been seven times greater still. And were there a pulsar in SN1987A, the ultraviolet and X-ray emission from its nebula would be expected to give highly ionized atoms in the surrounding gas, but emission from such ions has not been observed.

Even if the neutron star is rotating only slowly or has a weak magnetic field, one would expect surrounding material to fall back onto it, giving an Eddington luminosity of around $3 \times 10^{38} \text{ erg s}^{-1}$. The remaining possibility of a weak pulsar with little surrounding mass will be difficult to rule out. The thermal emission from a newly formed neutron star gives a luminosity of perhaps $3 \times 10^{34} \text{ erg s}^{-1}$, less than the expected power output from ^{44}Ti decays with a decay constant of 78 years. The neutron-star

emission should have a characteristic soft X-ray spectrum, but this emission will be absorbed by the surrounding supernova gas for hundreds of years.

If there is not a neutron star in SN1987A, could there be a black hole instead⁶? Had matter fallen back onto the nascent neutron star (with a mass of $1.4 M_{\odot}$) on a timescale of 100 seconds to a few hours after the explosion, enough mass may have been accreted to push the object over the minimum thought to be necessary for the creation of a black hole. Matter would continue to accrete onto the black hole, but the resulting radiation would be trapped in the inflow so that the escaping luminosity would be small. Because the cores of massive stars increase with the initial stellar mass, this picture would be more plausible if the initial mass of the SN1987A progenitor star, $20 M_{\odot}$, were close to the mass limit above which stars collapse directly to black holes. The putative black holes in Cygnus X-1 and LMC X-3 are thought to have formed from the collapse of massive stars and the lower limit for com-

plete collapse has generally been taken to be $40 M_{\odot}$. However, a new analysis by Maeder⁷, based on the amounts of helium and heavy elements ejected by massive stars, suggests that the lower limit is $20\text{--}25 M_{\odot}$. As the mass limit is increased, the yields of heavy elements strongly increase. So even if the astrophysical limits have deprived us of the opportunity to witness the emergence of a new pulsar, it may have given us a new black hole to look out for. □

Roger A. Chevalier is in the Department of Astronomy, University of Virginia, PO Box 3818, Charlottesville, Virginia 22903, USA.

1. Kurfess, J. D. *et al. Astrophys. J.* **399**, L137–L140 (1992).
2. Clayton, D. D. *et al. Astrophys. J.* **399**, L141–L144 (1992).
3. Woosley, S. E. & Hoffman, R. D. *Astrophys. J.* **368**, L31–L34 (1991).
4. Varani, G.-F. *et al. Mon. Not. R. Astr. Soc.* **245**, 570–576 (1990).
5. Suntzeff, N. B. *et al. Astrophys. J.* **384**, L33–L36 (1992).
6. Brown, G. E., Bruenn, S. W. & Wheeler, J. C. *Comments Astrophys.* **16**, 153–166 (1992).
7. Maeder, A. *Astr. Astrophys.* **264**, 105–120 (1992).

GEOCHEMISTRY

An element of recycling

Julian Pearce

ONE obvious and effective way of understanding the Earth's upper mantle is to study nodules (xenoliths) of mantle peridotite that are brought to the surface by volcanic eruptions. Unfortunately, the vast majority of mantle xenoliths are hosted by magmas that erupt well away from plate margins, few being found in lavas that erupt in volcanic arcs (where oceanic floor gets subducted) or at mid-ocean ridges (where it grows). This has not proved a great handicap for scientists studying mid-ocean ridges, as mantle is directly exposed at the sea floor in many transform faults and areas of strong amagmatic extension. It has, however, hampered work on volcanic arcs, where xenoliths present the only opportunity to sample the mantle directly underlying the volcanoes. On page 661 of this issue¹, Maury *et al.* present some of the first data from mantle xenoliths recovered from an arc volcano and help to answer an old argument on element recycling at subduction zones.

Subduction zones mark the principal sink for the return of the Earth's crust to the mantle from which it was once derived. In terms of mass balance, virtually all the subducted oceanic lithosphere is returned to the deep mantle where it is eventually reassimilated into the mantle convection system. In terms of chemical balance, however, the process is less straightforward. A substantial fraction

of the subducted mass of many trace elements is recycled before mixing with the deep mantle can take place. Some of this fraction is transported back to the surface in aqueous fluids released by dehydration reactions in subducted sediment and oceanic crust at shallow depths (see figure).

A further fraction is released by dehydration or melting of subducted materials at greater depths, where it promotes melting of the overlying mantle. In this case, the recycling medium is the magma erupted at the overlying volcanic arc, of which the recycled material forms a component. Such processes have a number of practical implications: for example, global geochemical budgets, including those of the greenhouse gases, cannot be fully balanced until element recycling at subduction zones is properly understood.

To date, some of the best information on element recycling has come from the trace-element signatures of volcanic-arc lavas. At ocean ridges or in most intra-plate settings, the concentrations of trace elements in volcanic rocks are controlled principally by the way these elements are partitioned between magma and the residual solid during melting. When trace-element abundances are normalized to an appropriate composition and plotted in order of affinity for the magma (incompatibility), smooth patterns emerge.

By contrast, lavas erupted above subduction zones typically have 'spiky' patterns, as element abundances are determined not just by their incompatibility but also by their ionic radius and ionic potential (charge/radius ratio). Typically, elements of low ionic potential, often termed the large-ion lithophile (LIL) elements, are enriched relative to those of high ionic potential, often termed the high-field-strength (HFS) elements. The effect, as the insets in the figure show, is that HFS elements such as Ta and Hf (also Nb and Zr) form negative anomalies with respect to surrounding LIL elements on geochemical patterns.

Without any knowledge of the mantle directly beneath the arc volcano, hypotheses for the origin of these anomalies have been difficult to prove. According to one hypothesis they are related to the transfer of elements from the subducted plate to the overlying mantle wedge. In this model, the LIL elements are more mobile than the HFS elements in the transporting aqueous fluid or melt. There is some experimental support for this idea. Tatsumi *et al.*² experimentally dehydrated serpentinite spiked with trace elements under subduction pressures and temperatures and found that the LIL elements were indeed more readily partitioned into the fluids released. However, there is a wide variety of subducted materials and not all have been studied in such detail.

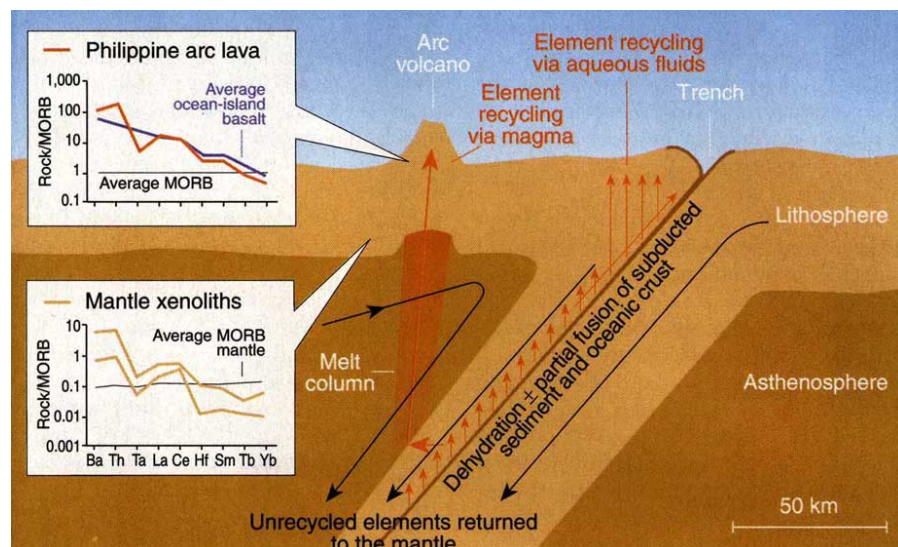
A second hypothesis is that titanate minerals are stable in the hydrous mantle beneath arc volcanoes and that these retain the HFS elements during the melting process. This model is out of vogue now that experimental work has shown that the minerals required to be stable do enter melt readily under these conditions³ but, again, not all the potential minerals may have been investigated.

Other hypotheses require that the two groups of elements are separated from each other during melt migration. Navon and Stolper⁴ argue that, just as elements in solution can be separated by flow through chromatographic columns in the laboratory, so the interaction of the upward-migrating magma beneath volcanic arcs with the surrounding mantle column could separate HFS and LIL elements. Keleman *et al.*⁵ propose that the same separation could be achieved by magma dissolving clinopyroxene and precipitating olivine on ascent. However, both models present the paradox that, despite their strong theoretical bases, they do not explain why the negative HFS element anomalies are largely restricted to volcanic eruptions above subduction zones.

Maury *et al.*, in this issue, demonstrate that their nodules from the Philippine arc have marked negative anomalies in

the HFS elements, especially Ta, similar to those found in arc lavas. This immediately confirms that the anomalies in the lavas do have a deep origin and cannot be explained by assimilation of crust. They also find highly variable enrichments of the LIL elements, those of highest radius (Cs, Rb, Ba) exhibiting the greatest enrichment. They contend that these enrichments are the result of interaction of mantle with an aqueous

been dredged and drilled in the inner walls of the deep ocean trenches. This mantle has a similar mineral composition to many of the Philippine nodules, but a very different geochemical signature. Zr and Hf, for example, exhibit positive anomalies with respect to adjacent LIL elements, the opposite of that observed in the Philippine nodules⁶. One explanation is that these signatures are caused by melting of the subducted basalt crust



Recycling of elements via aqueous fluids and melts at subduction zones. Maury *et al.*¹ find that both the lavas and the mantle xenoliths have negative anomalies in Ta content with respect to Th and La (insets). This anomaly could be caused by: (1) separation during dehydration and melting of the subducting plate; (2) a residual Ta-bearing mineral left after melting; or (3) separation in the melting column. It could also be caused by interaction between melt and crust. The xenolith data of Maury *et al.* favour the first hypothesis, that elements of large ionic radius are preferentially recycled at subduction zones.

fluid rather than melt, as melt-mantle interaction should also lead to enrichment (not found) in major elements such as Na and Al. The source of the aqueous fluid must, they argue, be the underlying subduction zone.

Significantly, concentrations of the HFS elements in the nodules are extremely low, lower than expected if residual titanate minerals were present or if the mantle column had selectively retained HFS elements. Proponents of the generation of HFS-element anomalies during melting or melt migration can argue with some justification that these processes may operate at greater depths than the 30-km source region for the xenoliths. Nonetheless, it is hard to dispute the contention of Maury *et al.* that their data best support the hypothesis that the anomalies in arcs can originate "from a combination of low initial concentrations [of the HFS elements] in the mantle . . . and the inability of hydrous [subduction] fluids to transport these elements". If true, subduction fluxes can be calculated on this basis.

Not surprisingly, the story does not end here. Mantle above subduction zones (though not from active arcs) has

when temperatures were hotter at the start of subduction.

Whatever the explanation, the Philippine nodules must be only part of the story. Any component driven off the subducted plate will vary according to the temperatures (and hence ages) of the subducting and overriding plates. It should also vary according to the nature of the crust and sediment being subducted. Clearly, precise assessments of global subduction fluxes are going to take some time. Nonetheless, the work by Maury *et al.* on the mantle nodules from the Philippine arc represents an important step in the right direction. □

Julian Pearce is in the Department of Geological Sciences, University of Durham, South Road, Durham DH1 3LE, UK.

1. Maury, R. C., Defant, M. J. & Joron, J.-L. *Nature* **360**, 661–663 (1992).
2. Tatsumi, Y., Hamilton, D. L. & Nesbitt, R. W. *J. volcanol. geotherm. Res.* **29**, 293–309 (1986).
3. Ryerson, F. J. & Watson, E. B. *Earth planet. Sci. Lett.* **86**, 225–239 (1987).
4. Navon, O. & Stolper, E. *J. Geol.* **95**, 285–307 (1987).
5. Keleman, P. B., Johnson, K. T. M., Kinzler, R. J. & Irving, A. J. *Nature* **345**, 521–524 (1990).
6. Parkinson, I. J., Pearce, J. A., Thirlwall, M. F. & Ingram, G. *Proc. ODP Sci. Results* **125**, 487–506 (1992).

Arterial hardening in mice

James Scott

By giving mice a gene that nature failed to, or, perhaps by a quirk of evolution, carelessly lost, R. M. Lawn and colleagues have managed to develop an animal model for atherosclerosis. The gene, for the blood protein apolipoprotein (a), is apparently missing from rodents and most other non-primate species (apart from the European hedgehog). When expressed in mice, as Lawn *et al.* describe on page 670 of this issue¹, the gene leads to the development of atherosclerosis in those mice on high-fat diets.

Atherosclerosis is a multifactorial disease of arterial hardening and fat deposition with both genetic and environmental components. The main risk factors are hypertension, dyslipoproteinaemia, coagulation disorders and cigarette smoking. However, with one exception (homozygous familial hypercholesterolaemia), there is no direct relationship between a single aetiological component and the development of the disease. Direct proof that an abnormality identified in population-based studies causes disease has been elusive, but is now rapidly becoming available as genes encoding key components for lipid transport in blood are either overexpressed in transgenic mice or knocked-out by homologous recombination².

Besides Lawn and colleagues' transgenic study, reports have appeared in recent issues of *Cell*³ and *Science*⁴ showing that knocking out expression of apolipoprotein E produces massive hypercholesterolaemia and profound atheroma even on a normal diet. And it seems that overexpression of apolipoprotein A-I⁵ and of the low-density lipoprotein receptor⁶ and its ligand apolipoprotein E⁷ (reported by E. Rubin at a recent meeting in Shizuoka, Japan) can protect susceptible mouse strains against diet-induced atherosclerosis.

In human population studies, lipoprotein (a) levels higher than 0.3 mg ml⁻¹, have been found in one in five individuals and are associated with a susceptibility to atherosclerosis, and therefore to myocardial infarction and stroke⁸. Lipoprotein (a) (Lp(a)) consists of a low-density lipoprotein (LDL)-like particle with a second protein; apolipoprotein (a) (apo(a)) tightly associated with the apolipoprotein apoB-100. Apo(a) is closely related to plasminogen, from which the enzyme that hydrolyses fibrin blood clots is released by tissue-plasminogen activator. The apo(a) molecule is composed of a variable repeat of plasminogen at its amino terminus, a disulphide-bonded kringle domain

shaped like a Danish pastry, known as kringle IV, a single kringle V, and the plasminogen protease domain. Apo(a) may be bonded to apoB-100 by a disulphide link, an association presumed to come about in the endoplasmic reticulum (ER), but we know from studies of patients with abetalipoproteinaemia — a rare mendelian recessive disorder in

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Deposition pattern of apo(a) in an atherosclerotic lesion of a human coronary artery. Immunohistochemical staining of the fibroatheroma demonstrates the accumulation of the apoprotein in bundle-like (open arrows) and punctuated (solid arrows) patterns, with the extracellular matrix. Staining for apoB-100 in a serial section shows almost identical patterns, revealing that the two are bound as Lp(a). (Micrograph courtesy of A. Niendorf and U. Beisiegel.)

which apoB-100 is synthesized but not secreted — that apo(a) can be secreted in the free form⁹.

In mice transgenic for human apo(a), this apolipoprotein is 95% free in the circulation¹⁰ (the baboon, too, secretes much of its apo(a) in the free form¹¹). When human LDL particles are infused into these transgenic mice they rapidly become tightly associated with apo(a) to form Lp(a) and can be dissociated only by reducing agents¹⁰. Although these experiments demonstrate the high affinity of apo(a) for apoB-100, they do not prove that a disulphide bond is formed, but human apo(a) probably cannot associate with murine apoB-100 because of the lack of conservation of cysteine residues at the carboxy terminus of rodent apoB-100^{12,13}. Normally, once apoB-100 (always in excess) and apo(a)

RESUME

SPOT the difference

PARTICIPANTS at last week's meeting of the American Geophysical Union, in San Francisco, were treated to the first ever satellite video of fault motion. R. L. Crippen of the Jet Propulsion Lab prepared the video from just a pair of pictures of the Mojave Desert in California where the Landers earthquake occurred earlier this year, one archival and one specifically requested after the rupture. Although the fault movement was less than a pixel, by flickering between the two frames, obtained with the French remote-sensing satellite SPOT, Crippen could discern the changed shadowing of the land form and even new fractures. The Global Positioning Satellite (GPS) system gives better spatial resolution, but it relies on sensors already put in place; Crippen's approach required no preparation.

Not so deadly

GENETIC engineering has given plants of deadly nightshade (*Atropa belladonna*) in which the alkaloid hyoscyamine, which this plant makes in abundance, is converted almost completely into the more commercially valuable alkaloid scopolamine, an anti-cholinergic drug used, for example, to prevent travel sickness. By inserting extra genes for the hydroxylase that converts hyoscyamine into scopolamine, and ensuring that they are expressed constitutively, D.-J. Yun *et al.* (*Proc. natn. Acad. Sci. U.S.A.* **89**, 11799–11802; 1992) produced plants containing most of the alkaloid content of leaves and stems as scopolamine. This also makes extracting the drug much easier; it can be simply crystallized out instead of being separated by chromatography from a complex mixture of alkaloids.

Flash of inspiration

OBSERVATIONS of γ -ray bursts by the Compton Gamma Ray Observatory have re-ignited a controversy: do the bursts come from our own Galaxy, or are they at cosmological distances? In principle, there is an easy test: if M31, the nearby Andromeda galaxy which closely resembles ours, has the same population of γ -ray bursts as our own Galaxy, then a sufficiently sensitive detector could spot them. Unfortunately, the Compton detector is not sensitive enough. But in the *Astrophysical Journal* (**400**, L59–L63; 1992) H. Li and E. P. Liang remark that X-ray flashes are sometimes associated with γ -ray bursts, with a typical relative output of a per cent or so. ROSAT, the X-ray satellite, is sensitive enough to see X-rays from M31, if all γ -ray bursts produce X-ray flashes at the 2 per cent level (which may not be the case). Several weeks' observation of M31 could then decide whether the bursts are galactic or not.

meet, they will become tightly associated. In abetalipoproteinaemia it may be that a microsomal triglyceride transfer protein which forms a dimer with protein disulphide isomerase, the ER enzyme responsible for the correct arrangement of disulphide bridges, may be absent¹⁴. These results are consistent with a model for the assembly of triglyceride-rich VLDL (very-low-density lipoprotein) particles and Lp(a) in which disulphide formation and lipidation are co-translational events. Do apoB-100 and apo(a) associate at this early stage or later in the secretory pathway?

The transgenic mice of Lawn and colleagues¹ show for the first time that free circulating apo(a) can lead to atheroma, fat deposits, on a high-fat diet. The lesions in the mouse, a species not normally susceptible to atheroma, closely resemble the human disease in consisting of lipid-rich fibrous plaques. This begs the questions of how the atherosclerosis is caused, and what the normal function of Lp(a) is.

It may be the Lp(a) adds an extra control to fibrinolysis by competing for the binding of plasminogen to its endothelial cell receptor, thereby blocking the activity of tissue plasminogen activator, when in excess it would induce a procoagulant state⁸. However, Lp(a) binds avidly to tissue matrix components, including glycosaminoglycans, fibronectin, collagen and fibrin.

The presence of apo(a) with apoB-100 in atherosclerotic plaques (see figure) and the binding of kringle IV to fibrin has led to the suggestion that Lp(a) delivers cholesterol to proliferating cells at sites of vascular injury. Excess local Lp(a) at injury sites may be oxidized and taken up by macrophages, and contribute to foam-cell formation within plaques¹⁵. The present work provides little support for this cholesterol-delivery hypothesis, rather supporting the view that apo(a), independent of its lipid component, perhaps inhibits clot lysis and that this can initiate the activation and accumulation of lipid-laden macrophages and smooth muscle cells.

Apo(a) shows extraordinary size variation, assuming a relative molecular mass between 280,000 and 830,000. More than twenty size alleles are inherited in a co-dominant fashion and result from variation in the number of kringle IV repeats in the gene. The plasma concentration of Lp(a) is inversely correlated with allele size. Ninety per cent of the variability in Lp(a) concentration in the plasma is attributable to the Lp(a) locus. Much, but not all of this is due to size variation¹⁶.

Until now, the promoter of the gene has eluded identification. The reason is that apo(a) and plasminogen and all other closely related genes reside on

chromosome 6. Complementary DNA probes from the 5' end of apo(a) DNA reveal at least six distinct genomic fragments showing more than 90 per cent nucleotide homology over more than a kilobase. Three of these have been identified as belonging to plasminogen, and are termed plasminogen-related genes A and B¹⁷; the latter is expressed in tumours¹⁸. The other three sequences may be part of apo(a)-like genes, and/or pseudogenes produced by duplication. At the meeting in Shizuoka, Lawn reported that a region 1 kilobase long from the 5' flanking regions of all six of these loci can direct the transcription of a luciferase reporter gene in HepG2 cells, so that the products attributable to these promoters and the effect of transcription on Lp(a) blood levels should soon be discovered.

The identification of the apo(a) gene promoter(s) has another importance. There is no effective treatment yet for reducing elevated blood Lp(a) concentration. The known hypolipidaemic drugs (HMG-CoA reductase inhibitors, bile acid sequestrants, fibric acid derivatives, probucol) and diet have little effect on Lp(a) concentrations. Nicotinic acid alone or in combination with neomycin can reduce Lp(a), but unpleasant flushing and hepatic and ototoxicity militate against the long-term use of these drugs. However we are entering a new era in which gene promoters attached to marker genes are being screened in order to discover compounds that interfere with transcription. This is made possible by the development of advanced robotics and computing, and recombinant DNA methods for generating huge libraries of compounds. Lipoprotein(a), apoA-I, the microsomal triglyceride transfer protein and their genes are prime targets for the pharmaceutical industry. □

James Scott is in the Department of Medicine, Royal Postgraduate Medical School, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK.

1. Lawn, R. M., Wade, D. P., Hammer, R. E., Verstuyft, J. G. & Rubin, E. M. *Nature* **360**, 670–672 (1992).
2. Brown, M. S. & Goldstein, J. L. *Cell* **71**, 187 (1992).
3. Plump, A. S. et al. *Cell* **71**, 343–353 (1992).
4. Zhang, S. H., Reddick, R. L., Piedrahita, J. A. & Maeda, N. *Science* **258**, 468–471 (1992).
5. Rubin, E. et al. *Nature* **353**, 265–267 (1991).
6. Hofmann, S. L. et al. *Science* **239**, 1277–1281 (1988).
7. Shimano, H. et al. *Proc. Natn. Acad. Sci. U.S.A.* **89**, 1750–1754 (1992).
8. Scott, J. *Br. Med. J.* **303**, 663–664 (1991).
9. Talmud, P. J. et al. *J. clin. Invest.* **82**, 1803 (1988).
10. Chiesa, G. et al. *J. biol. Chem.* (in the press).
11. White, A. L., Rainwater, D. L. & Lanford, R. E. *J. Lipid Res.* (in the press).
12. Law, A. & Scott, J. *J. Lipid Res.* **31**, 1109 (1990).
13. Yang, C.-Y. et al. *Arteriosclerosis* **9**, 96–108 (1989).
14. Wetterau, J. R. et al. *Science* **258**, 999–1001 (1992).
15. Zionscheck, T. F., Powell, L. M., Rice, G. C., Eaton, D. L. & Lawn, R. M. *J. clin. Invest.* **87**, 767–771 (1991).
16. Boerwinkle, E. et al. *J. clin. Invest.* **90**, 52–60 (1992).
17. Ichinose, A. *Biochemistry* **31**, 3113–3118 (1992).
18. Weissbach, L. & Treadwell, B. *Biochim. biophys. Res. Commun.* **186**, 1108–1114 (1992).

Travelling light

DAEDALUS once invented a steam lawnmower that burnt its own cuttings. The longer and thicker the grass, the more copious its fuel supply and the more vigorously it worked; and when its task was completed it stopped automatically. He now links this idea to his plan of last week for an agricultural diesel engine that burns powdered plant waste. It carries a cryogenic system to pulverize the waste at liquid-air temperature, and to condense its exhaust to a slurry of plant ash, water and nitrates of combustion. This is returned to the soil as fertilizer. Daedalus sees this principle as the basis of a whole new ecologically benign transport system.

Imagine, he says, a road covered not with concrete but with grass. It would probably need a concrete meshwork of some sort to support vehicle wheels, but most of its area would be under cultivation. A vehicle travelling along the road could shave a small amount from the grass, and burn it for fuel. By the time the next vehicle came along, the grass would have grown again.

Such a road could power light traffic indefinitely, at no cost. If the grass were to trap solar energy with a photosynthetic efficiency as great as 10 per cent, then in full sunlight the agricultural road could fuel about one vehicle every ten seconds. In practice, a vehicle every few minutes might be the upper limit. This could be increased by wide traffic-lanes, enabling each vehicle to harvest a different track of grass. The system is even better suited for powering a light railway. Railway traffic rates are much lower than those of a road, and the energy needed to carry a given weight at speed is much less.

Grass is quite a good plant for this job. It can be flattened by a wheel and springs back undismayed; it survives repeated cropping by lawnmowers or cows; it can hold its own against weeds, and will dominate a mixed, equilibrium ecosystem. DREADCO biologists would prefer a plant with a higher photosynthetic efficiency, but the best performers in this category seem to be algae like *Chlorella* which flourish in water. So Daedalus also plans to adapt his scheme for transport on a system of deliberately eutrophic canals, possibly the receivers in a distributed sewage system. Always green, and scummy with algae, such canals could power a steady stream of boats that sucked in and burnt the floating plant material.

This elegant arrangement is clearly unsuited to intensively industrial societies. Daedalus offers it as a contribution to the eco-dream of a sustainable, postindustrial society, such as may develop in the Third World when the oil has run out. David Jones

Cause of wild dog deaths

SIR — African wild dogs (*Lycaon pictus*) are highly endangered, and their conservation is a top priority of the IUCN canid specialist group¹. The wild dog population of Serengeti National Park has recently undergone a drastic, disease-driven decline. Burrows² has suggested that the population's demise is attributable to activation of latent rabies infections as a result of the stress of handling (darting, anaesthesia, radiocollaring or vaccination). He hypothesizes that handling may stimulate the production of adrenocorticosteroids, leading to immune suppression and thus to increased vulnerability to disease.

Because post-vaccination immunosuppression can lead to clinical distemper in captive African wild dogs³, the hypothesis warrants careful testing. Furthermore, the hypothesis calls into question fundamental methods of studying wildlife; if typical handling is capable of inducing immune suppression, substantial changes in methods would be needed for many studies. In this light, several points are relevant.

Citing ref. 4, Burrows states: "handling-induced stress, as measured by highly elevated peripheral serum cortisol concentrations, results from immobilization of captive wild dogs". Four features of these data should be noted. (1) Cortisol was measured with a human cortisol radioimmunoassay kit⁴, with no validation for *Lycaon*. (2) The wild dogs in this study were darted and anaesthetized 13 times in 51 days. As a result, they showed panic when darting was attempted, running directly into chain-link fences⁴. This is in sharp contrast to anaesthesia in the field, where darted wild dogs react little. (3) After handling, mean serum cortisol levels were 181 nmol l⁻¹ ($n=35$). Baseline plasma cortisol levels are not known for *Lycaon*, but assuming a baseline similar to that of domestic dogs, repeated darting caused short-term cortisol peaks that were 1.4-fold above baseline; this is a mild elevation in comparison to corticosteroid peaks in other species. For example, in wild dwarf mongooses (*Helogale parvula*) captured in livetraps, urinary free cortisol increased 6.8-fold above baseline ($t_{242} = 14.15$, $P < 0.001$; S. C. and S. Monfort, manuscript in preparation).

(4) Most important, corticosteroid immune suppression results from chronically elevated adrenocorticoid levels, and not from adaptive short-term increases. No data allow us to address whether handling induces a long term stress response in *Lycaon*. Indeed, no previously published data have tested for chronic corticosteroid elevations in response to typical handling in the field,

for any species. In Serengeti dwarf mongooses, trapping, anaesthesia and handling induced a significant short-term increase in urinary free cortisol. But in the month following trapping, mean urinary cortisol levels dropped to 10.1 ± 4.9 ng per mg creatine (mean $\pm$ s.e., $n = 23$), below the baseline (27.2 ± 1.8 ng per mg creatinine, $n = 234$). Thus, for the only species for which data are available, handling does not induce chronic corticosteroid elevation. For wild dogs, we have no data to address the question, and an answer awaits endocrine data from the field (now under way).

Are there alternative explanations to be considered? Wild dogs are susceptible to several fatal diseases, including distemper, rabies and anthrax^{2,5}. The history of the Serengeti population shows that this disease outbreak is the latest in a series. In 1967–68, the population was not handled, yet three packs were noted to have been more than halved by disease (probably distemper)⁶. In 1971–73, the population was not handled, but the population fell from 96 to 49, with 5 of 12 packs disappearing; again, disease was implicated⁷.

Studies of four major wild dog populations so far show no effect of darting and radiocollaring on mortality (vaccination has not occurred in these populations). 'Handling' is common to stable populations and the rollercoaster Serengeti population. Other factors differ, such as Serengeti's history of low population density and repeated bottlenecks⁷. These factors could be involved in the recent crash, and should be given equal consideration.

Scott Creel

Wildlife Conservation Research Unit,
Department of Zoology,
Oxford University, Oxford OX1 3PS, UK

1. Ginsberg, J. R. & Macdonald, D. W. *Foxes, Wolves, Jackals and Dogs: An action plan for the conservation of canids* (IUCN, Gland, Switzerland, 1990).
2. Burrows, R. *Nature* **359**, 277 (1992).
3. Durchfield, B., Baumgartner, W., Herbst, W. & Brahm, R. *J. vet. Med. Ser. B37*, 203–212 (1990).
4. van Heerden, J. & Kuhn, F. S. *Afr. J. Wildl. Res.* **15**, 80–84 (1985).
5. van Heerden, J. S. *Afr. J. Wildl. Res.* **16**, 7–11 (1986).
6. Schaller, G. B. *The Serengeti Lion* (Univ. Chicago Press, 1972).
7. Malcolm, J. R. thesis (Harvard Univ. 1979).

SIR — Burrows's Scientific Correspondence draws attention to the controversy over the possible connection between rabies vaccination and associated handling of the African wild dog, *Lycaon pictus*, and their catastrophic mortality in 1990–91 in the Serengeti/Masai Mara ecosystem¹. *Lycaon* is a highly endangered, flagship species numbering fewer than 5,000 individuals distributed in fragmented populations². Any con-

cerns regarding the effectiveness of research on, or conservation of, this species demand close scrutiny.

Burrows raises misgivings about the need for and results of the vaccination programme. His simple question, "why vaccinate?", has a complex answer. Rabies characteristically causes high mortality in most canids^{3,4}. Many countries in eastern Africa, including Kenya, have record numbers of confirmed cases, mainly in dogs⁵. Rabies has been confirmed in Serengeti bat-eared foxes (B. Maas, personal communication), and a reportedly rabid hyaena in the Mara has recently killed a child. Rabies was confirmed in five *Lycaon* individuals in the Serengeti/Mara ecosystem; the virus isolated from one of these is similar to the strain found in domestic dogs in Africa. Blood samples were drawn from 12 *Lycaon* in the Serengeti, three of which showed >0.5 IU per ml of rabies serum-neutralizing antibodies (RSNA)⁶. The measured presence of antibodies in unvaccinated animals could be caused by non-specific neutralizing substances, but titres of >0.5 IU per ml in domestic dogs are generally considered specific to prior exposure to rabies virus (M. Fekadu, personal communication). These antibodies do not necessarily mean that an animal is protected; it may be incubating the disease. But because all except one of the sampled *Lycaon* were seen 5 months later it is more likely that those with >0.5 IU per ml RSNA had survived earlier exposure. Protection from natural rabies virus infection requires both humoral (antibodies) and cell-mediated (activated lymphocytes) responses; it does not follow from the presence of humoral antibody that an animal is protected. Even if some individuals were immune, near-total population immunity to rabies has never been reported for any canid. For all these reasons, and because rabies (and distemper) was observed in domestic dogs within and adjacent to the ecosystem (K. Alexander, personal communication), rabies is clearly a disease risk for *Lycaon*, and one with potentially devastating consequences.

Why did rabies vaccination fail? The most plausible reason for so many *Lycaon* dying despite vaccination against rabies is that they did not die of rabies. The data are too few to determine if rabies caused the deaths in *Lycaon* up to 8 months after vaccination, but if it did either they were incubating the disease before vaccination or vaccination failed to protect them against subsequent challenge. The incubation period is unknown in *Lycaon*, but may exceed 12 months in domestic dogs⁷ and in some wild African carnivores it may exceed 6 months (J. Bingham, personal communication). Vaccination could fail if: (1) the vaccine

had not been inactivated; (2) the dose was insufficient; (3) the immunity induced by vaccination was overwhelmed; (4) procedural inadequacies compromised the vaccine; or (5) colostral antibody interfered with the normal immune response in young animals. Differential survival of vaccinated dogs in the Mara (1/19 unvaccinated; 1/3 vaccinated) suggests the vaccine had some efficacy (P. Kat and J. Richardson, personal communication). In other species, similar doses have provided protection for 5 years or longer⁸.

Could vaccination and handling contribute to the demise of *Lycaon*? The safety (but not the efficacy) of inactivated rabies vaccines for this species has been tested twice. Madivak was administered to *Lycaon* in Frankfurt Zoo; two *Lycaon* were vaccinated with Rabisin in the Mara in 1987 and four in 1989 (J. Richardson, personal communication). No ill effects were observed in either case. Inactivated rabies vaccines would not give rise to clinical rabies, but could accelerate the deaths of individuals already incubating rabies⁹.

Although post-vaccination immunosuppression can lead to clinical canine distemper in captive *Lycaon*¹⁰, Burrows's hypothesis that stress due to handling may reactivate latent rabies virus in carrier *Lycaon* is speculative. There is scant evidence that handling is dangerously stressful (see Creel's letter above), that latency exists in rabies^{7,11}, that stress reactivates virus in survivors of non-lethal rabies in any mammal, that stress accelerates incubation in any wild canid (but see ref. 11), or that a carrier state exists in wild canids (but see refs 12, 13). We conclude that while stress might accelerate the onset of clinical signs in individuals incubating rabies, it is extremely unlikely that stress reactivated rabies in individuals that had previously experienced a non-lethal infection.

Burrows's data show increased mortality in 1990–91, concurrent with handling and bounce vaccination. With the

fragmentary data available we cannot conclusively reject any hypothesis, but we consider that the handling-induced-rabies hypothesis is improbable. A simpler hypothesis compatible with these data is that longevity of *Lycaon* was negatively correlated with year, irrespective of handling, because mortality pressure has increased due to other factors, for example distemper². The Serengeti data are insufficient to test such hypotheses, so we recommend that further handling of the Serengeti *Lycaon* is suspended pending immediate analysis of the survival of handled *Lycaon*. Serious screening in the Serengeti ecosystem for other directly transmitted diseases of Carnivora should be fostered, and monoclonal antibody and other techniques be used to characterize local strains of rabies virus. Vaccine efficacy should be explored using *Lycaon* held in captivity and unsuited for release or captive breeding programmes. All this work should be coordinated among ecologists and wildlife pathologists.

Burrows constructively draws attention to the catastrophic decline of this

species and provokes questions about the effectiveness of the research done on *Lycaon* in the Serengeti/Mara. Although the decision to vaccinate *Lycaon* seems reasonable to us, it is regrettable that the subsequent debate has been fermented by the inadequacy of data on the fates of handled *Lycaon* in the ecosystem. Of course it is easy to be wise with hindsight, and money is often a limiting factor, but the conservation of this marvellous species in an ecosystem it has helped make famous deserves the highest standards of scientific application.

D. W. Macdonald*

Chairman, IUCN/SSC Canid Specialist Group,
Wildlife Conservation Research Unit,
Department of Zoology,
Oxford University,
Oxford OX1 3PS, UK

* Other signatories to this letter: M. Artois, M. Aubert, CNEVA, BP 9, 54200 Malzeville, France; D. L. Bishop, NERC Institute of Virology, Oxford OX1 3SR, UK; J. R. Ginsberg, Institute of Zoology, London NW1 4RY, UK; A. King, Central Veterinary Laboratory, Weybridge KT15 3NB, UK; N. Kock, University of Zimbabwe, PO Box MP 167, Mount Pleasant, Harare, Zimbabwe; B. D. Perry, International Laboratory for Research on Animal Diseases (ILRAD), PO Box 30709, Nairobi, Kenya.

Ancient microbe database

SIR — Occasionally one reads of anecdotal references about researchers who have discovered microorganisms that have been preserved for thousands of years. Some researchers have even revived and isolated these organisms. Examples include the isolation of organisms from spores upon opening a tomb of a long-buried Pharaoh, obtaining organisms from the gut of a preserved mastodon, digging up bottles of beer from sunken galleons, or obtaining cultures from rocks or other sedimentary material. These accounts raise several opportunities. For example, these re-

vived or preserved organisms offer a unique opportunity to study evolutionary genetics. A snapshot of the past is captured for study today. They may also have desirable industrial characteristics that have since been lost by their contemporary descendants. Already a porter beer has been commercialized from a 167-year-old porter yeast.

These accounts also raise several issues. Should these isolates be treated with special care because they are potentially different to the gene pool existing in nature today? Is it possible that an extinct disease-causing microorganism

ORGANISMS PRESERVED FOR MORE THAN 100 YEARS

NONViable Organism	Length of preservation (years)	Isolated from	Ref.
Iron bacteria	2×10 ⁹	Rocks from Western Australia	1
Boring green alga	408–421×10 ⁶	Sedimentary rocks of Eastern Poland	2
Fomes species spores	4,300–4,900	Rocks from lake site excavations in France	3
Caccidioides immitis	600–1,000	Skeleton of an Indian unearthed at an archaeological site in Northern Arizona	4
VIABLE			
Bacillus circulans	650×10 ⁶	Salt deposits	5
Enterobacter cloacae	11,000	Intestinal tract of a mastodon	6
E. agglomerans	11,000	Intestinal tract of a mastodon	6
Thermoactinomyces species	1890	Roman archaeological site	7
Porter yeast	167	1825 shipwreck	8

- Burrows, R. *Nature* **359**, 277 (1992).
- Ginsberg, J. R. & Macdonald, D. W. *IUCN World Conservation Union* (Gland, Switzerland, 1990).
- Macdonald, D. W. & Voigt, D. R. in *Population Dynamics of Rabies in Wildlife* (ed. Bacon, P. J.) 71–108 (Academic, London, 1985).
- Baer, G. M. (ed.) *The Natural History of Rabies* (2nd edn) (CRC, Boca Raton, 1991).
- Perry, B. D. in *Rabies in Eastern and Southern Africa* (in the press).
- Gascoyne, S. C. et al. *J. Wildlife Dis.* (in the press).
- Fekadu, M. in *The Natural History of Rabies* (2nd edn) 192–198 (ed. Baer, G. M.) (CRC, Boca Raton, 1991).
- Aubert, M. F. A. *Rev. Sci. Tech. Off. Int. Epiz.* (in the press).
- Brochier, B. M. et al. *J. gen. Virol.* **70**, 1601–1604 (1989).
- Durchfield, B. et al. *J. vet. Med. Ser. B37*, 203–212 (1990).
- Blancou, A. et al. *Rabies inf. Exch.* **3**, 7–9 (1980).
- Fekadu, M., Shaddock, J. H. & Baer, G. M. *Am. J. trop. Med. Hyg.* **30**, 1113–1115 (1981).
- Kovalev, N. A. & Sedov, V. A. *Conf. Commission regionale de l'OIE pour l'Europe Rap. no. 205* (1971).

could be brought back to life with disastrous effects? If microorganisms have been revived, what precautions have been taken to establish that the revived organism is indeed revived and not simply an environmental contaminant?

To begin grappling with some of these issues and to present a hitherto under-used resource, we have set up a database of organisms which have been preserved for more than 100 years. This database is split into two sections. The first contains information about organisms that, although dead, have been preserved. DNA material may or may not be available from these organisms. The second contains a list of those organisms that have been revived after storage for long periods of time (see table).

The database is currently in its infancy, but it is hoped that others may contribute information about other long preserved microorganisms so that a central list can be kept as a potential resource. Researchers knowing of any other long-preserved organisms are encouraged to contact us directly.

Max J. Kennedy

Sarah L. Reader

The New Zealand Institute for

*Industrial Research and Development,
PO Box 31-310, Lower Hutt, New Zealand
Fax: 64 4 569 0132*

1. Karkhanis, S. N. *Nature* **261**, 406–407 (1976).
2. Kazmierczak, J. & Golubik, S. *Nature* **261**, 404–406 (1976).
3. Purdy, L. H. & Purdy, B. A. *Bot. Gaz.* **143** (4), 551–553 (1982).
4. *Anon Science* **253**, 1353 (1991).
5. Kendig, F. & Hutton, R. *Life-Spans or How Long Things Last* (Holt, Rinehart and Winston, New York, 1979).
6. Park, P. *New Scientist*, p 12 (1 June 1991).
7. Seawa, M. R. D., Cross, T. & Uns, B. A. *Nature* **261**, 407–408 (1976).
8. Coghlan, A. *New Scientist*, p 2 (14 September 1991).

Familiar strangers

SIR — All three protein structures^{1–3} reported in a recent issue of *Nature* display folding motifs previously observed in other protein structures. In only one case did the authors concerned notice this similarity.

The amino-terminal domain in the staphylococcal enterotoxin B (SEB) structure¹ seems to be the fifth example of a novel common fold. This fold consists of a five-stranded β -barrel capped by an α -helix between the third and fourth strands. It has been observed in three other proteins secreted by bacterial pathogens: staphylococcal nuclease⁴, the B-subunits of heat-labile enterotoxin⁵ and verotoxin-1 (ref 6). It has also been found in the anticodon-binding domain of Asp-tRNA synthetase⁷. These four proteins, with very different amino-acid sequences, are very similar in secondary structure, even in minor details. They all bind oligonucleotides or oligosaccharides and their binding sites, where

known, are similarly located on a particular side of the barrel⁸. These observations suggest that the corresponding site on SEB's amino-terminal domain may be essential for its toxic function. The carboxy-terminal domain of this protein seems to be a variation of the β -grasp motif⁹ previously observed in ubiquitin, chloroplast-type ferredoxins and in the immunoglobulin G-binding domain of streptococcal protein G.

Musacchio *et al.*² found no protein fold closely related to the one they revealed in the structure of Src-homology (SH3) domain of spectrin. But this fold has been observed previously in the structure of R67 dihydrofolate reductase¹⁰. The fold, common to both structures, is composed of an antiparallel β -sheet of five strands coiled into a barrel-like structure. The common secondary structure also includes a turn of 3₁₀ helix between the fourth and fifth strands. Despite having different amino-acid sequences, the close structural resemblance of SH3 and R67 proteins may be a clue to an as yet unidentified function of the SH3 domain.

The histidine-rich actin-binding protein hisactophilin was described in ref. 3 as having the same fold, the β -trefoil fold¹¹, as interleukin-1 and fibroblast growth factors. It is actually the eleventh example of this fold, which has also been observed in Kunitz trypsin inhibitors, ricin B chain¹¹ and in the interleukin-1 receptor antagonist protein¹². The amino-acid sequences of proteins from different functional families show few significant similarities. However, in addition to their similar fold the proteins share another common feature: they all perform recognition functions.

Not only these three, but a substantial fraction of proteins whose structures have now been determined and which have no clear sequence similarity to previously known structures, seem to share already known folding motifs⁸. This information supports the suggestion that a few different protein folds perform a much larger number of biological functions.

Alexey G. Murzin

*MRC Laboratory of Molecular Biology,
Hills Road, Cambridge CB2 2QH, UK*

1. Swaminathan, S. *et al.* *Nature* **359**, 801–806 (1992).
2. Musacchio, A. *et al.* *Nature* **359**, 851–855 (1992).
3. Habazettl, J. *et al.* *Nature* **359**, 855–858 (1992).
4. Arnone, A. *et al.* *J. biol. Chem.* **246**, 2302–2316 (1971).
5. Sixma, T. K. *et al.* *Nature* **351**, 371–377 (1991).
6. Stein, P. E., Boodhoo, A., Tyrrell, G. J., Brunton, J. L. & Read, R. J. *Nature* **355**, 748–750 (1992).
7. Moras, D. *Trends biochem. Sci.* **17**, 159–164 (1992).
8. Murzin, A. G. & Chothia, C. *Curr. Opin. struct. Biol.* **2**, 895–903 (1992).
9. Overington, J. P. *Curr. Opin. struct. Biol.* **2**, 349–401 (1992).
10. Matthews, D. A. *et al.* *Biochemistry* **25**, 4194–4204 (1986).
11. Murzin, A. G., Lesk, A. M. & Chothia, C. *J. molec. Biol.* **223**, 531–543 (1992).
12. Stockman, B. J. *et al.* *Biochemistry* **31**, 5237–5245 (1992).

DNA correlations

SIR — Despite your editorial¹ saying that the long-range correlations in DNA reported by Peng *et al.*² and by Voss³ had yet to be explained, we had already proposed one explanation of the phenomenon^{4,5}.

We concluded that non-coding DNA sequences (introns) tend to have longer-ranged correlation than coding DNA sequences (exons). Some, but not all, intron sequences exhibit even longer-ranged correlation and the resulting power spectra are $1/f^\alpha$ ($\alpha \approx 1$)^{4,5}. Both conclusions appeared in refs 2 and 3, confirming our findings. We had proposed that the longer-ranged correlation is due to the abundance of repetitive patterns which are common in intron sequences.

We were led to this study when one of us was searching for long-range correlation and $1/f^\alpha$ spectra in cellular automata — a class of sequence manipulation rules which do not change the sequence length. It was soon realized that it is much easier to generate long-range correlation when the sequence length is allowed to increase.

This observation led to the study of a class of mathematical models called expansion-modification systems in which the sequence lengths become longer and longer^{6,7}. These systems are extremely simple, and the only two basic operations contained are duplication (or amplification if more than two copies are made) and mutation. It was found that if the two operations are applied repeatedly, the elongated sequences can have $1/f^\alpha$ power spectra.

According to common belief, the lengths of present-day DNA sequences are longer than those of three billion years ago, when life started. One of the mechanisms for sequence elongation is gene or oligonucleotide duplication⁸, during which copies of a segment of the sequence are made and these copies are inserted back into the original sequence. Similar to the case in expansion-modification systems, the gene or oligonucleotide duplication process is not perfect. If it occurred at different stages of evolution, the length of the repeated pattern could be different, leading to correlations at different length scales.

The fingerprints of oligonucleotide duplication are preserved better in intron sequences where, as we have indeed

1. Maddox, J. *Nature* **358**, 103 (1992).
2. Peng, C.-K. *et al.* *Nature* **356**, 168–170 (1992).
3. Voss, R. *Phys. Rev. Lett.* **68**, 3805–3808 (1992).
4. Li, W. & Kaneko, K. *Europhys. Lett.* **17**, 655–660 (1992).
5. Li, W. *Int. J. Bifurc. Chaos* **2**(1), 137–154 (1992).
6. Li, W. *Europhys. Lett.* **10**, 395–400 (1989).
7. Li, W. *Phys. Rev.* **A43**, 5240–5260 (1991).
8. Ohno, S. *Evolution by Gene Duplication* (Springer, Heidelberg, 1970).

observed, the long-range correlation is more prominent. In Voss's paper, the only two categories (phage and bacteria) where the white-noise component dominates contain no introns, consistent with our explanation that long-range correlations are perhaps due to the repetitive patterns that are observed most often in introns.

Wentian Li

Cold Spring Harbor Laboratory,
PO Box 100,

Cold Spring Harbor, New York 11724, USA

Kunihiko Kaneko

Department of Pure and Applied Science
University of Tokyo

Komaba, Meguro

Tokyo 153 Japan

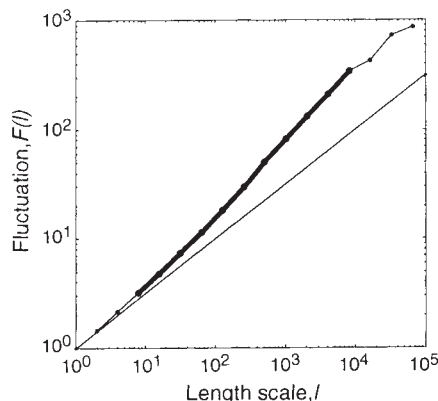
SIR — Peng *et al.*¹ reported apparent long-range correlations within various intron-containing gene sequences up to length scales of about 10 kb, but not in intronless genes. Voss² later reported such correlations, manifested as spectral density estimates that obey the $1/f^\beta$ law, where $\beta \sim 0.8$, in DNA sequences from various organisms. Our analysis of the first completely sequenced eukaryotic chromosome³ (*Saccharomyces cerevisiae* chromosome III, 315 kb), unlike that of Prabhu and Claverie⁴, supports those findings and extends the length scales over which correlations are found. Moreover, we provide a simple statistical test for the significance of such correlations based on the analysis of variance.

The complete yeast chromosome III sequence (EMBL accession number X59720) encompasses roughly 182 putative coding regions, of which only 3% contain introns. It also contains substantial stretches of presumably noncoding DNA in gene-flanking regions. We have not concerned ourselves here with the question of identifying intronless sequences (see ref. 4), but ask whether there are indeed long-range correlations within the entire chromosome.

Following the method of Peng *et al.*, we have plotted the fluctuation $F(l)$, that is, the root mean square (r.m.s.) of the differences between the purine and pyrimidine frequencies in all regions of length l , for $l = 2, 4, 8$ and so on. We find that the fluctuation grows as l^α , with $\alpha \sim 0.7$ when the length scale l is greater than 10 (see figure). (Although α here is related to β above, there is no explicit connection between them, and the terminology we use follows that in the references cited herein.) The slope of this plot (α) remains statistically significantly greater than 0.5 (the expectation for an uncorrelated series) up to about $l = 8$ kb. Such higher than expected slopes indicate correlations in the purine/pyrimidine ratio in regions of length l . The longest sequence analysed in ref. 1, the 73-kb human β -globin region,

showed a significantly elevated slope only up to about $l = 1$ kb, although the apparent high slope extends further.

The statistical significance can be assessed by analogy to standard statistical analysis of variance. Each point in the figure is obtained by first dividing the series into $2m$ non-overlapping regions of length l and computing the fluctuation, $F(l)$. By joining the neighbouring pairs of such regions, we create m new regions of length $2l$, and compute the fluctuation for these longer regions, $F(2l)$. The purine/pyrimidine differences in the new, larger regions have a fluctuation or variance related to the 'between sum-of-squares' of the analysis of variance, whereas the fluctuation in the original regions is related to the 'total sum of squares'. The statistical significance of any excess variance in the larger



Fluctuation of the purine-pyrimidine random walk in yeast chromosome III plotted against length scale. Intervals with significantly larger than expected slope ($P < 0.01$) are shown with a heavy line. The straight line has the expected slope, $\alpha = 0.5$, for an uncorrelated series.

regions (the main-effect) can be obtained by comparing

$$\frac{(m-1) F^2(2l)/2}{(2m-1) F^2(l) - (m-1) F^2(2l)/2}$$

to the critical value of Fisher's F -distribution with m and $m-1$ degrees of freedom.

The correlations we observe could result from long regions of shifted base composition. As Nee⁵ shows, such regions can result in slopes greater than 0.5 on the fluctuation plot. However, unlike Nee, we believe that the observations of Peng *et al.* do imply a lack of independence of the purine/pyrimidine occurrences, at long ranges. Because the purines and pyrimidines tend to be loosely clustered in the DNA sequence, knowledge that a particular base within a cluster is a purine augments the probability that another, possibly distant base in the same cluster is also a purine. That is, the bases within a cluster or region are not acting independently. Such regions are well known. For example, CpG

islands up to ~ 150 bases in length, or multiple repetitive elements arranged in clusters extending up to about 10 kb are found in other genomes. Thus, some patterns have already been observed on length scales similar to that of our finding. Voss found long-range correlations in the 229 kb cytomegalovirus genome (our test gives significance out to $l = 16$ kb), but the biology of viral genomes is known to differ markedly from that of nuclear chromosomes (for example, the replication cycle does not require interaction with a centromere for segregation). Our finding of significant correlation in a complete, functional chromosome at scales eight times larger than in the β -globin gene region supports the idea that such correlations are a general property of eukaryotic DNA, and not necessarily limited to isolated examples. Repeating the analysis, counting the number of weakly pairing (A or T) versus strongly pairing (C or G) nucleotides or counting any single nucleotide (A, T, C or G), showed similar results.

It will be interesting to see if statistically significant correlations are observed at still larger scales (100 kb, 1 Mb), when still longer DNA sequences are determined. Although the mechanisms giving rise to these long-range patterns or correlations remain unexplained, such mechanisms are possibly involved in the functioning, regulation and perhaps even of evolution of the genome.

Peter J. Munson

Ronald C. Taylor

George S. Michaels

Division of Computer Research and
Technology,

National Institutes of Health,

Bethesda, Maryland 20892, USA

1. Peng, C.-K. *et al.* *Nature* **356**, 168–170 (1992).
2. Voss, R. F. *Phys. Rev. Lett.* **68**, 3805–3808 (1992).
3. Oliver, S. G. *et al.* *Nature* **357**, 38–46 (1992).
4. Prabhu, V. V. & Claverie, J.-M. *Nature* **357**, 782 (1992).
5. Nee, S. *Nature* **357**, 450 (1992).

Gas correction

SIR — Perfect gases (see J. Maddox *Nature* **359**, 669; 1992) are modelled as comprising point particles having mass but not volume, billiard-ball or otherwise. If they evinced volume, they would be imperfect, and instead of $pV = nRT$, we should have to replace V by, for example, the van der Waals ($V - nb$), where b represents the sum of the particle volumes in a mole of the gas; or by more elaborate devices, as listed by Maddox.

David R. Rosseinsky

Department of Chemistry,
The University,
Exeter EX4 4QD, UK

Literal interpretations

George M. Marsden

The Creationists: The Evolution of Scientific Creationism. By Ronald L. Numbers. Knopf: 1992. Pp. 456. \$27.50.

IMAGINE yourself deeply committed to a community that confirms your experience that its religious faith is the most important commitment that a person can have. Integral to this community is belief in an unerring Bible. Moreover, your community's faith is energized by the conviction that biblical prophecies will be fulfilled precisely by events accompanying the imminent return of Jesus. The whole Bible is correspondingly to be interpreted as literally as possible. Hence the Earth cannot be much more than 6,000 years old. A web of reinforcing factors makes it virtually impossible for you to abandon any of the above beliefs. Accepting an old Earth or biological evolution is therefore out of the question. Your job, as someone fascinated by science, is to use your high intelligence to find an alternative model in which to fit the scientific evidence. The task is difficult, but no more so than other challenges that scientists have surmounted.

Ronald Numbers, a distinguished historian of science at the University of Wisconsin, understands well the drama and the potential agony of the challenge facing creation scientists. Numbers recounts that he grew up in a family of Seventh-Day Adventist preachers but that he eventually came to the devastating conclusion that attempts to explain the geological evidence by a world-wide flood was a hopeless task. Nonetheless, in this monumental history of the young-Earth movement, there is no bitterness or cynicism. Rather, he tells the story with remarkable even-handedness. This feat reflects the unique perspective of one who is no longer an Adventist, yet who keeps on his desk a framed ticket for a 1940s lecture entitled "God's Answer to Evolution", in which his father was the lecturer.

This marvellously detailed, engagingly told and sometimes astonishing history has the intrigue of a Chaim Potok novel. Like Potok's *The Chosen*, in which an Orthodox Jewish boy finds out that his orthodoxy is liberal in the view of his Hasidic neighbours, *The Creationists* is largely the story of a struggle between the orthodox and the hyperorthodox. One of the remarkable parts of the story is how a small group of flood geologists rose from a marginal position, even among conservative Bible-believing Protestants, to come to represent the best-known creationist viewpoint.

Before 1960, what is today known as 'creation science' had only the most meagre support even among the conservative evangelical or fundamentalist communities in the United States. Most earlier fundamentalist leaders had allowed for some accommodation of the geological evidence for an old Earth. Some allowed that the "days" in Genesis might represent aeons. Many others subscribed to a "gap theory" that allowed for vast amounts of time between when

were broadening their outlooks and beginning to call themselves 'evangelicals'. The American Scientific Affiliation, founded in 1941 as an organization of Bible-believing scientists, originally included proponents of a young Earth as well as advocates of an old Earth. During the next two decades, this sizeable organization became a forum for accommodations of biological evolution and biblical belief. The small minority of young-Earth proponents felt excluded and resolved to establish an alternative for fundamentalists that would be equally viable scientifically.

Central in the ensuing transformation of flood geology into an influential national movement was Henry Morris. Morris, who is described by Numbers as a person of ability and integrity, was a fundamentalist fascinated by prospects



Making waves — Henry Morris (left) and John Whitcomb in 1984. Their influential book *The Genesis Flood* (1961) was a classic of creation science.

the Universe was created in Genesis 1:1 and its being "without form and void" in Genesis 1:2. Even William Jennings Bryan, leader of the antievolution crusade after the First World War, allowed for an old Earth. Just before the Second World War, almost the only prominent figure who insisted on a young Earth and argued that the biblical flood provided a scientific explanation of geological data was George McCready Price, a tireless lecturer dedicated to vindicating Adventist founder Ellen White's revelation on these points.

During the 1920s, however, most American fundamentalists had become militantly opposed to biological evolution, especially in reaction to those who were using Darwinism to shock people with faith in a literal Bible. Between 1940 and 1960, however, opposition to biological evolution was weakening among fundamentalist academics who

for literal fulfilments of biblical prophecies. Unhappy with views of the Bible that did not take it at face value, he dedicated himself to defending a young Earth, adopting flood geology much like Price's. Receiving a PhD in hydraulic engineering from the University of Minnesota in 1950, he held an important position in engineering at Virginia Polytechnic Institute until 1969, when he felt forced out because of his fundamentalist views. In the meantime, the landmark for the launching of his alternative movement was the publication with John Whitcomb, a biblical scholar, of *The Genesis Flood* (1961), a volume that eventually sold more than 200,000 copies. The book helped to spark the formation in 1963 of the Creation Research Society under Morris's leadership.

Not happy with the narrow sound of the term 'flood geology', Morris referred

From *The Creationists*

to his movement as "creation science". Throughout the 1960s, creation scientists emphasized frankly the biblical basis for their views. During the 1970s, however, the strategy shifted. One of the explanations for the surge in popularity of the movement was that by the 1960s biological evolution had been reintroduced into US public schools and was seen by fundamentalists as one part of a new relativizing of American public values. In the midst of the conservative backlash of the 1970s, creation scientists attempted to insert their views into public schools as an alternative to evolutionary views. This public strategy necessitated shifting emphasis to the scientific, rather than the biblical, basis of their conclusions. Furthermore, their characteristic view of science shifted from a Baconian objectivism to appeals to the views of Karl Popper and Thomas Kuhn, which emphasized that equally valid science may proceed from alternative viewpoints.

In order to win legislative approval, flood geologists now claimed for themselves the term 'creationists', ignoring the fact that every other type of Christian believed in some form of creation as well. To the biblical literalists, however, all views that said that God used evolutionary processes as a means of creation were inconsistent. Therefore, they argued, there were only two real options: evolutionism and 'creationism', meaning flood geology. Incredibly, in 1980 the legislatures of Arkansas and Louisiana adopted this 'two-model' view, mandating the teaching of 'creationist' (in Arkansas specifically young-Earth) models alongside evolutionary views. Although these laws were struck down in the courts, the creation-science movement has continued to flourish at the local level. Moreover, polls show that nearly half the US population will affirm recent creation of humans. Creation-science arguments have also been exported world-wide to new churches where biblicist 'either/or' arguments have strong appeal.

Numbers does a marvellous job of telling this story, particularly the internal history of the creation-science movement. His research is superb, bringing in much previously uncited correspondence. Even when he is describing some of the questionable activities of some of the more marginal figures who have been part of the movement, his tone is even-handed, letting the facts speak largely for themselves.

Numbers does not spend a great deal of time analysing the reasons for the astonishing success of the movement. One factor he mentions is the populist appeal. American evangelicalism has long had a democratic rhetoric, appealing to people to decide for themselves.

In American folk religion there is also a tremendous reverence for the Bible, literally interpreted. Creation scientists have been able to validate their arguments and to win legislative support by appealing to this popular base. The



From *The Creationists*

Price: tireless anti-evolutionist.

movement also has other political connections, which Numbers only mentions. Creation science typically has been closely associated with a broader conservative religious-political package. Morris, for instance, has long been a

friend of Jerry Falwell. Numbers is, however, clear on the general point that creation science is a reactive movement. As in Bryan's time, the perception that evolutionary naturalism is undermining all traditional values is one impetus for promoting stark alternatives.

One lesson implicit in this history is that some scientists too have been guilty of posing stark alternatives. Evolution has often been used to ridicule any traditional faith. Some secularists have been all too ready to accept flood geologists' claims to speak for all 'creationists' and then to dismiss even more nuanced arguments that belief in a creator might be a useful hypothesis for understanding the Universe. Those who insist that nothing could ever be clarified by positing the existence of a higher intelligence are in a formal sense something like creation scientists. They are so committed to a community that finds this secular faith immensely useful that they are convinced that they *must* be able to find an exclusively naturalistic explanation for everything.

Unless we suppose that natural science of the past century has somehow settled all such issues, we might expect that in the future there would be wider acceptance of a variety of hypotheses about the possible relationships of natural phenomena to higher creative intelligence. Those who insist on the extremes, however, may delay any such evolution of scientific thought. □

George M. Marsden is in the Department of History, University of Notre Dame, Notre Dame, Illinois 46556, USA.

Chance and ad-hackery

Betty A. Toole

Accidental Empires: How the Boys of Silicon Valley Make Their Millions, Battle Foreign Competition, and Still Can't Get a Date. By Robert X. Cringely. Addison-Wesley/Viking: 1992. Pp. 336. \$19.95, £16.99.

ROBERT X. Cringely is the pseudonym of the author of this very readable and entertaining book about the history of the modern computer industry and the personalities behind it. The mask evokes as many rumours as the book itself, because Cringely writes the gossip column for *Infoworld*, which is read by about half a million people. His column has been compared to Kitty Kelley, *Private Eye* and *Spy* — those fascinating collections of half-truths that many of us devour.

Like many other light and humorous works, this book's leitmotif is in fact

serious. In proclaiming in the preface that there are no morals in the book, Cringely is both right on target and wide of the mark — "no morals" is an oxymoron. Of course, there *is* a moral, found in the title *Accidental Empires*: namely, that good guys don't always come first. This comes as a welcome relief to all those popular science books that leave the reader with the depressing feeling that people successful in science and technology are somehow deserving. In reality, quality does not necessarily guarantee success. This is certainly true for the rise of the personal computer.

After mentioning the first stage of the Silicon Valley revolution, masterminded by the "old boys" such as Robert Noyce and Gordon Moore, Cringely focuses on the development of the personal computer, the software designed for it, and the "boys" such as Bill Gates, Steve Jobs and Mitch Kapor who drive the

industry. The juicy stories are about Gates, the Rajneesh of the computer world. They range from metaprogramming, to shipping software filled with bugs, to his queuing for some ice cream while fumbling for a 50-cent discount coupon, holding everyone up until someone eventually gives him the 50 cents (which he takes without a pause). Variations of the ice-cream story, and others like it, have become the litany of the computer world. But this does not distract from the prescience of Gates, who organized Microsoft into a business that works, even if the company's software often enters the market filled with bugs (the third versions usually work).

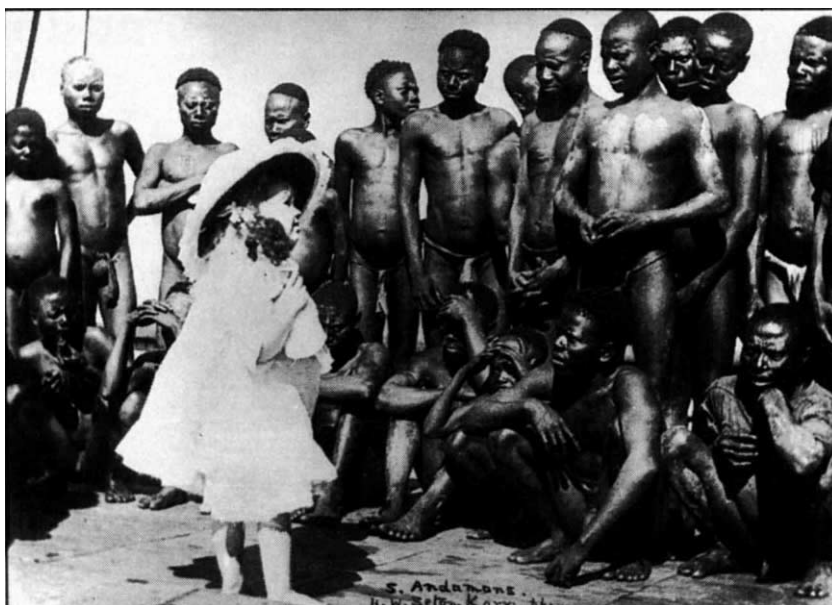
Cringely also vividly describes the recent, and to him continued, decline of IBM. IBM feared the competition that the personal computer presented to its mainframe business and as a result underestimated the growing power and popularity of the personal computer. There is a need for a new standard, however: many of the players today will not be there in the future because, as Cringely emphasizes, standards are based more on accident and increasing returns than on quality. He writes:

To Engineers — really good ones, interested in making progress — the best of all possible worlds would be one in which technologies competed continuously and only the best ones survived . . . [But] the real world, the one we live in, is a world of dollars, not sense. It's a world where commercial interests are entrenched and consumers pay more attention to what everyone else is buying than to whether what they are buying is any good.

Cringely does not believe that the future belongs to IBM or to Japanese manufacturers. He recognizes an important shift now in personal computing, citing four notable trends: the growth of standard-based computing, RISC (reduced instruction set computing) processors and advanced semiconductors, and the death of the mainframe. These trends will create a wealth of opportunities in the development of software and microprocessors. Whether the paradigm of accidental empires continues to characterize the nature of the next shift will be fascinating to see. I, for one, cannot help but wonder what the title of Cringely's sequel will be. □

Betty A. Toole is a consultant for Critical Connection, PO Box 452, Sausalito, California 94966, USA.

■ Computer experts, high-tech millionaires and teenage whiz-kids are also a few of the subjects of Bruce Sterling's *The Hacker Crackdown: Law and Disorder on the Electronic Frontier*, which claims to be the first full look at computer hacking. Bantam/Viking, \$22.50, £16.99.



Ethnic minorities — Mary Deane, daughter of the census officer, on board a government steamer with a group of Onges, Little Andaman, 1911. The picture is one of 157 remarkable photographs reproduced in *Anthropology and Photography, 1860–1920* edited by Elizabeth Edwards. Yale University Press, \$35, £19.50.

Black holism

John D. Barrow

Black Holes. By Jean-Pierre Luminet. Cambridge University Press: 1992. Pp. 312. £30, \$59.95 (hbk); £10.95, \$22.95 (pbk).

BLACK holes are the simplest things in nature. Attempt to give a complete description of something even as mundane as this page and you would be faced with amassing countless pieces of information about atomic numbers, bindings and so on, together with all manner of large-scale properties such as temperature and size. But find yourself a black hole and there are only three things about it that you can know: its mass, its spin and its electric charge. It is not that the material within the horizon of the hole has no other properties or becomes transformed into some mysterious anonymous form unlike the stuff of the world around us; indeed, nothing at all unusual happens to it on entering the hole. It is simply that these three properties of the totality of material encompassed by the space-time boundary of the black hole are the only pieces of information about the interior that are accessible to outside observers.

This elegant minimalism is one reason for the attraction of black holes to gravitation physicists. But astronomers are enticed by the apparent inevitability of black holes as the final resting places of sufficiently massive stars that have ex-

hausted their reserves of nuclear fuel, and by the seeming ubiquity of black holes in the most interesting places in the Universe. The strong gravitational field gradients surrounding them can liberate quantities of radiant energy far in excess of the energy produced by any other known force. As a result, they are prime candidates to explain a plethora of phenomena in the realm of high-energy astrophysics. There is compelling evidence that black holes lie at the heart of the quasar phenomenon, lurk in the centres of active galaxies, and orbit in tandem with luminous stars whose fragile outer layers are dragged down into the whirlpool of the black hole amidst a flux of tell-tale X-rays.

There have been many books about black holes at a popular low level and there are several mathematical treatises for specialists. In between, there is something of a vacuum. Filling it is a challenge, because black holes are studied in a variety of ways by specialists interested in quite different aspects: some workers are attracted by the clues that these objects give to the intrinsic properties of gravity; others build intricate computer simulations of the collapse of material that leads to their formation; still others seek to model complicated astronomical observations by combinations of black holes, stars and gas. The wide range of expertise required to bring all this together is perhaps why it is rare to encounter a single unified exposition of black-hole properties.

Yet Jean-Pierre Luminet has managed

to achieve this synthesis in brilliant style. He is a researcher who is active in the study of many astrophysical features of black holes and his striking simulations of the optical appearance of black holes are presented here in greater detail than ever before (but still without the use of mathematics). His style is lively and engaging, enlivened by interconnections with a wider culture than the purely scientific, and the translators have done an excellent job in reproducing some of the linguistic elegance of the original French. The amount of information conveyed is impressive. The intended audience is readers with some understanding of physics who are seeking a coherent, accurate nonmathematical overview of black-hole physics and all the astrophysical situations that the discipline seems to explain. Also, any student embarking on a serious technical study of general relativity or astrophysics will find the book a first-rate overview of an important part of the story. The treatment is broad and balanced, and Luminet is enviably patient in his determination to explain as much as possible and to make black holes as simple as possible, but no simpler. This is an outstanding work of scientific exposition that can be recommended as a manifestation of the dictum that common sense often leads to surprises and the modest goal of science is simply to make such surprises just a little less frequent. □

John D. Barrow is in the Astronomy Centre, University of Sussex, Brighton BN1 9QH, UK.

Bench notes

An unusual laboratory diary to regulate the life of the harassed biochemist is on offer in the form of *The 1993 Axel Biochemical Year Planner*. Instead of containing unit-conversion tables and information on women's shoe sizes, this A5 week-to-view diary is crammed with lists of biomedical conferences to be held during 1993, more than 400 international research institutions, some 100 pharmaceutical companies, and databanks of cell and animal collections (with telephone and fax numbers). At the back are compilations of roughly 250 common biochemical abbreviations; data on molecular weights, isotopes, restriction enzymes, DNA and proteins, and antibiotics and galactosides; laboratory recipes; acid, base and buffer charts; and even a sort of "Freezer Stock Location" filofax. And sitting rather self-consciously among all this is a wine chart — a fitting seasonal entry for the nonspecialist. Published by Axel Press Ltd, Kennett House, 108/110 London Road, Oxford OX3 9AW, UK. Price £3.99 (perfect bound), £6.99 (spiral bound). **R.C.**

Onco-promotion

John Cairns

Genes and the Biology of Cancer. By Harold Varmus and Robert A. Weinberg. *Scientific American Library* (W. H. Freeman): 1992. Pp. 224. \$32.95, £17.95.

TEN years ago, two papers appeared in *Nature* announcing that the tumorigenicity of a human cancer was due to a point mutation in an oncogene. Since then, information has poured forth about the sequence changes in cancer cells. The normal functions of the altered genes are now fairly well understood, and the particular sequence changes in the different types of cancer have been shown, in some cases at least, to match what is known about the particular mutagenicity of the causative agent. The story is a perfect example of how molecular biology has clarified biology and biochemistry. At last, cancer starts to be understandable.

Varmus and Weinberg give a clear and exciting account of these developments, in which they themselves played a major part. The very beginning of the story was somewhat before their time, and this is the only part of their account that I would question. They say: "By the late 1940s . . . the target of carcinogens was no longer elusive". In fact, the cancer establishment reacted extraordinarily slowly to the coming of molecular biology. E. Boyland, P. Brookes and P. Lawley were about the only people in cancer research in the 1940s and '50s who believed that chemical carcinogens interact with DNA. In 1952, Boyland was ridiculed in print by E. and J. Miller who felt sure that the target was protein. As late as 1963, H. Pitot and C. Heidelberger were arguing that carcinogens interact with the proteins that regulate gene expression. What settled the matter were the experiments of Brookes and Lawley in 1964 and the results of microbial mutagenicity assays using liver extracts to achieve metabolic activation — a test invented by F. Mukai, W. Troll and H. Malling in 1969 and then systematized by Bruce Ames. But these are minor details. The book is excellent, and I am sure it will be very useful to all those biologists who want the latest news about the oncogenes.

I do not think the book will be of much help to the general public, but it will make a handsome addition to the art books on coffee tables. And that may have been exactly what the publishers had in mind when they created the *Scientific American Library* Series. For example, the first set of pictures in *Genes and the Biology of Cancer* shows a fertilized egg, a blastula, a mammalian

embryo and the final product — a young girl about to plunge her chopsticks into what appears to be a bowl of bouillabaisse. Those four pictures (in exquisite colour) take up as much space as the subsequent brief introduction to DNA, RNA, protein and the genetic code. No reader who needs to be told about cells will understand such a compressed summary.

Scientists have a difficult time when they try to explain their work to the lay public. The safest bet is to go for a strictly historical approach as used by D. Dressler and H. Potter in the companion volume called *Discovering Enzymes*. But it is not easy. Begin right at the beginning, and you will feel that you are wasting too much time on the preliminaries; begin *in medias res*, and you know you will be incomprehensible. So you aim for something in between and rely on your nonscientific friends and your copy editor to tell you where you are taking too much for granted. That, I suspect, was the stage missed out in *Genes and the Biology of Cancer*. It is, for example, fine to have a picture illustrating the transformation of pneumococci by DNA, but you must tell the reader that it shows the two types of bacterial colonies, not the bacteria themselves; without this explanation, the photograph is a waste of space because it will be understood only by those who have already seen it. The same can be said for many of the diagrams; it is common practice to show DNA sequences as a series of boxes (representing the exons, introns, long terminal repeats, insertion sequences, promoters, operators and so on), but this kind of abstraction should be preceded by a clear description of the relationship between sequence and function.

If the target of the book is the general public, I think that much more space should have been devoted to the origins and ideas of molecular biology; after all, these are what underlie all the things now known about "genes and the biology of cancer". If the target is primarily undergraduates and senior high-school students (all of whom already know about cells, if not about DNA), they should not have to pay for the many unnecessary colour pictures, put in simply to guide the book onto a few coffee tables. □

John Cairns is at the Clinical Trials Service Unit, Harkness Building, Radcliffe Infirmary, Oxford OX2 6HE, UK.

■ *Scientific American Books* (W. H. Freeman) recently published a revised second edition of the acclaimed introductory text *Recombinant DNA* by J. D. Watson, M. Gilman, J. A. Witkowski and M. J. Zoller. £33.95 (hbk), £21.95 (pbk).

Evolution and environment in the Hominoidea

Peter Andrews

Between 10 and 20 million years ago, a variety of hominoid primates lived in Africa, Europe and Asia. The question of which of these, if any, lie closest to the ancestries of humans and modern apes remains a lively source of debate. Recent fossil discoveries, though, shed light on the environments in which the various groups of hominoid emerged and, it is hoped, on their evolution. But the lack of a hominid fossil record before about 5 million years ago—and any fossil record for the African apes—is still a frustrating barrier.

THE hominoid primates are the group that encompasses the apes and humans. The origin of the group and the evolution of the various hominoid species are of relevance to the human condition, but of more immediate concern are the issues of when and where the human species originated and the context in which it evolved: the effects of climate, environment and behaviour, for example, and the ecological relationships between communities. I shall review the evidence relating to these issues, particularly the systematic relationships of fossil hominoids and their palaeoecology, insofar as it is currently available from the fossil record.

The hominoid apes include the lesser apes, the gibbons and siamangs of eastern Asia, and the great apes, chimpanzees and gorillas of Africa and the orang-utan of South East Asia. Molecular evidence has shown that the old concept of great apes (family Pongidae), as distinct from humans (family Hominidae), is no longer valid¹ because the African apes and humans are more closely related to each other than any of them are to the orang-utan. This is recognized taxonomically by the division between the Ponginae, which contains only the orang-utan and its putative fossil relatives, and the Homininae, for the African apes and humans.

This distinction is supported both by molecular evidence and by morphology. Neither source of evidence, however, offers a definitive solution on the relationships within the African ape

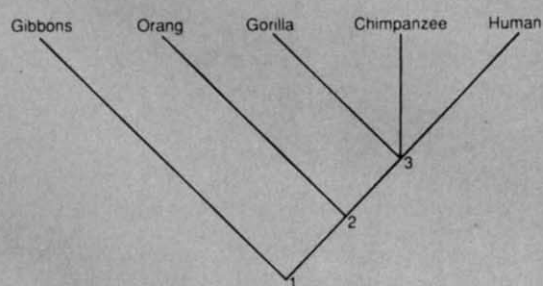
and human group, the Homininae. The majority of molecular studies support a closer relationship between chimpanzees and humans, with gorillas more distantly related (see Box 1), but some molecular studies support the alternative of closer relationship between chimpanzees and gorillas, with humans the sister group to the combined African apes. Morphological evidence is also ambiguous, so in my opinion the trichotomy shown in Box 1 is not yet resolved.

Hominoid origins

The earliest fossil record for the Hominoidea is in early Miocene deposits in East Africa. The genus *Proconsul* is widely distributed in Uganda and Kenya in sediments ranging in age from 22 to 17 Myr (million years ago)². *Proconsul* is recognized as having mainly ancestral catarrhine characters of the skull and dentition, but to be linked with the Hominoidea on the basis of a few postcranial characters^{3–5}. Some of these have been appropriately criticized⁶, but I consider that the characters listed in Box 2 are well established as hominoid synapomorphies shared by *Proconsul*. It has also been claimed recently that *Proconsul* lacked a tail⁷, as do extant hominoids, and that it had a relatively larger brain than comparably sized monkeys⁸ (Box 2), and there is now little doubt that *Proconsul* is phylogenetically linked with the Hominoidea.

BOX 1 The relationships of the extant hominoid primates

The characters distinguishing the three branching points are listed below.



Node 1 Characters of the extant Hominoidea.

Large numbers of characters defining the ancestral hominoid morphotype have been listed^{3,6,75,76}, but the most useful way of summarizing them is to extend Harrison's⁶ list of discrete functional complexes for the postcrania:

1. differential usage of the forelimb, including increased potential for raising arms above the head, for extending the forelimb at the elbow joint, and for rotation of the forelimb;
2. greater flexibility of the wrist and opposable thumb;
3. more erect posture during locomotion and feeding with broadening of thorax and loss of tail;
4. greater mobility at the hip and ankle joints.
5. To this may be added the characters of the dentition related to diet,

such as broad spatulate central incisors, low crowned premolars with loss of honing on P3, and relatively broad molars with low rounded cusps.

Node 2 Characters of the Hominidae linking Ponginae and Homininae^{3,6,76}.

1. Skull with enlarged maxillary sinuses, orbits higher than broad, increased alveolar prognathism with elongated premaxilla, facial lengthening with elongation of nasal bones and narrow incisive foramen;
2. great increase in size of incisors relative to molar size;
3. robust and enlarged premolars relative to molars, upper premolar heteromorphy reduced;
4. reduced molar cingula;
5. mandible robust with large inferior transverse torus;
6. distal humerus with deep sulci either side of lateral trochlear keel.

Node 3 Characters of the Homininae linking the African apes and humans. The characters distinguishing this node are both morphological (described in more detail in Box 3) and molecular^{76,77}, for example:

1. myoglobin chain, positions 23 and 110;
2. two shared substitutions in the fibrinopeptide A and B chains;
3. twenty-seven shared substitutions in the DNA and mitochondrial DNA chains and three deletions;
4. fusion of os centrale in the wrist;
5. enlarged supra-orbital torus.

Evidence resolving the trichotomy at node 3 is provided for a human-chimpanzee grouping^{78–87} and for a chimpanzee-gorilla grouping^{88–96}. Some of these workers recognize that the trichotomy is not yet resolved^{97–101}, but others insist that it is resolved in favour of the chimpanzee-human grouping^{81,87,102}.

Species diversity in the early Miocene East African hominoids is high, with two to three species of *Proconsul*-related forms and two to three species of small-bodied apes present at many sites⁹. Differences are partly based on size: the two species of *Proconsul* from Rusinga Island are estimated at 9 kg and 26–38 kg based on good postcranial evidence¹⁰. The small species, formerly *P. africanus*, is now being referred to a new species¹¹, with *P. africanus* being restricted to Koru and Songhor, together with the similarly sized but morphologically distinct *Rangwapithecus gordonii*. Also present at these sites is a much smaller form related to the latter and a larger species of *Proconsul*, *P. major*.

Associated with *Proconsul* are a number of less well known fossil primate taxa, many of which are inferred as being hominoid but which lack the necessary body parts to be sure. The 'small-bodied ape' *Limnopithecus legetet* has been grouped with *Proconsul* in the Hominoidea on the basis of dental evidence¹². An older discovery that has become significant is the taxon "*Proconsul* (*Xenopithecus*) *hamiltoni*" described for a maxilla fragment from Lothidok Hill in northern Kenya¹³, because the Eragaleit beds from which this fossil came (together with some additional undescribed specimens) have recently been dated at between 24.3 and 27.5 Myr¹⁴. This is several million years older than the previously oldest known fossil ape from Meswa Bridge in the Koru area¹⁵, and if this specimen is indeed a fossil hominoid it puts the date for hominoid origins at least to ~25 Myr.

Proconsul has been shown to have been mainly associated with forested palaeoenvironments in East Africa. The faunas and floras from Songhor, Koru and Mwangano in Kenya, and Napak in Uganda all have clear affinities with present day rainforest biotas¹⁶. Some of the Rusinga Island floras are closer to more seasonal tropical woodlands, although there are wet forest elements as well¹⁷, and the faunas have both forest elements and non-forest at different stratigraphic horizons¹⁶,

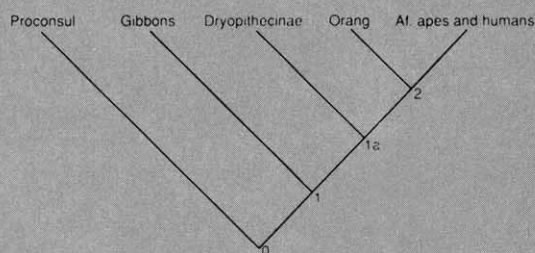
indicating a mixture of habitats during the Miocene. The community structure of the mammalian faunas is similar to modern forest communities¹⁸, again with the exception of some of the Rusinga faunas. It may be that the species difference between the Rusinga hominoids and those from all the other East African sites reflects these apparent ecological differences, for there are similar differences in several other groups of mammals¹⁸.

All of the *Proconsul* species appear to share a form of arboreal quadrupedalism¹⁹ that was little changed from their catarrhine ancestors²⁰. Evidence from the wear patterns of teeth shows that they also shared similar diets of soft fruit and young leaves²¹, and it seems likely that the various species operated within similar ecological constraints, although there is evidence of some ecological diversification in contemporaneous hominoids both in diet, for example the more folivorous *Rangwapithecus*²¹, and in locomotion, for example the long-limbed *Dendropithecus*²². The recent conclusion that the *Proconsul* hand had a fully opposable thumb, with rotation at the trapezium-first metacarpal joint at least as extensive as in other hominoids²³, is of particular significance in combination with thumb length, for *Proconsul* lacks the relatively short thumb of the living great apes and so had hand proportions very similar to those of modern humans²⁴. This is particularly interesting in view of the recent emphasis on tool use in chimpanzees²⁵, and it would be interesting in this respect to ascertain if there is any evidence of tool use by natural living populations of gorillas and orang-utans.

The Dryopithecinae

In the period from 17 to 12 Myr there appears a second radiation of hominoid species. Their relationships with each other and with later hominoids is still uncertain, but as they represent different evolutionary trends I am going to distinguish them taxonomically at the level of tribe, where they have been distinguished at generic level before. Three groups of fossil hominoid

BOX 2 Relationships of the fossil hominoids, with characters distinguishing the branching points



Two fossil clades are recognized here. The first consists of the Proconsulidae at node 0 and includes *Proconsul*, for which the best evidence is available. This clade also includes *Rangwapithecus*, *Nyanzapithecus*, and the Lothidok hominoid (see text). Some of the characters distinguishing node 0 are the same as for node 1, thereby providing evidence for the hominoid status of *Proconsul*, as follows:

1. low crowned P3 with length greater than height, cusp heteromorphy on upper premolars reduced, reduction in breadth of upper molars (breadth/length 115–121%), and I¹ nearly as broad as high (breadth/length index 80–90%)¹²;
2. relative increase in brain size: encephalization quotient 48.8% compared with a range of values of 22 to 41% for eleven species of monkey⁸;
3. medial torsion of the humeral head and elongated vertebral border of the scapula, leading to increased potential for raising arms above the head⁵;
4. strong medial and lateral keels of the trochlea²⁴;
5. increased mobility at the ulnocarpal joint¹⁰³;
7. phalanges of hallux broad and robust, and complete opposability of the thumb²³;
8. absence of tail⁷.

The second node, at position 1A on the cladogram, is probably a heterogeneous association of taxa, grouped here because we do not at

present understand their phylogenetic relationships. It is referred here to the Dryopithecinae and it includes three groups: (1) *Dryopithecus* itself in tribe Dryopithecini; (2) a newly proposed tribe, Afropithecini with *Afropithecus*²⁶, *Heliopithecus*³² and *Otaviopithecus*^{33,34}, together with material from Maboko Island and Nachola that has been referred in the past to *Kenyapithecus*; and (3) the Kenyapithecini, which includes *Kenyapithecus wickeri* from Fort Ternan³⁸, *Griphopithecus alpani* from Pasalar, Turkey³⁷, and the postcrania from Klein Hadersdorf³⁶. The characters distinguishing node 1A are as follows:

1. premolar enlargement combined with retention of varying degrees of cusp heteromorphy³²;
2. molars with low cusps, flattened occlusal surfaces⁴⁷, reduction in cingulum and squarish shape³⁴;
3. reduction in size of superior transverse torus of the mandibular symphysis³;
4. robust but low crowned canines.

At node 1A, the three groups have the following distinctive characters:

Afropithecini. Increase in enamel thickness of the molars to intermediate thick, with relative enamel thickness⁴⁸ of 17.35 (ref. 47); and further enlargement of the premolars. The postcrania (from Maboko⁴²) retain primitive hominoid characters similar to the morphology seen in *Proconsul*.

Kenyapithecini. Increase in enamel thickness, but the enamel is even thicker than in afropithecins, as measured for the Pasalar sample, with relative enamel thickness of 19.71 (ref. 47), although it has yet to be measured for any of the African *Kenyapithecus*; this character may be diagnostic of node 1A if it can be shown to be ancestral for both pongines and hominines⁴⁸;

Dryopithecini. Limb bones with rounded shafts; distal humerus with rounded capitulum, deep olecranon fossa and prominent trochlear keels⁴⁹. These characters may be homologous with those of the orang-utan, or with those of the African apes and humans, but not with both as there is now evidence of independent derivation of orang postcranial characters from the other great apes⁵⁴.

may be recognized, manifesting two rather distinct evolutionary trends. The three groups, and their constituent fossils, are listed in Box 2.

The fossil record of the first group, the Afropithecini, is confined to the African continent. Three important fossil sites are known, the Turkana sites of Buluk and Kalodirr yielding *Afropithecus*, dated to older than 17 Myr²⁶, the Arabian site of Ad Dabtiyah, yielding *Heliopithecus*, with a faunal age of about 17 Myr²⁷, and the East African sites of Maboko Island and Nachola, yielding '*Kenyapithecus*' *africanus*, and both dated about 15 Myr^{28,29}. These three genera seem to form a natural grouping (Box 2) and in the past they have been confused with one another: in commenting on the original description of the Buluk specimens, Deslon³⁰ suggested that they should be attributed to *Kenyapithecus* (rather than to *Sivapithecus*, as they then were³¹), and in an addendum to their original description of *Heliopithecus*, Andrews and Martin³² recognized the affinities of that genus to *Afropithecus*, which was published just before *Heliopithecus*. The recently described taxon *Otaviopithecus namibiensis* from 13 Myr deposits in Namibia³³ has been compared with this group³⁴ on the basis of shared dental characters, but cranial and postcranial material is needed to clarify this point. This combination of specimens could be recognized either by grouping them into a single genus or into a tribe, and on present evidence I favour the latter alternative and suggest the Afropithecini as a suitable name. This follows from the probable distinction between the Fort Ternan and Maboko^{35,36}/Nachola²⁹ material: because the type species of *Kenyapithecus* is *K. wickeri* from Fort Ternan, the generic name *Kenyapithecus* has to remain with the Fort Ternan material and forms the root for the second tribe Kenyapithecini.

Later in the Miocene, the hominoid fossil record becomes sparse in Africa. *Kenyapithecus* is known from middle Miocene sites such as Fort Ternan³⁵ (see above), about 14 Myr, but the rest of the fossil record is restricted to a small number of indeterminate isolated teeth and fragmentary hominine jaws (see below). Similar aged sites in Eurasia are also rare, but hominoids very similar to *Kenyapithecus wickeri* are known from 14 to 15 Myr sites in Turkey, Pasalar and Candir³⁷, and Czechoslovakia and Austria (Neudorf Sandberg and Klein Hadersdorf). These fossils constitute the second group recognized here, the Kenyapithecini³⁸. The Pasalar hominoids have been assigned in the past to *Sivapithecus*, but new material shows that at least one of the species lacks the derived subnasal morphology of that genus. For this reason, we have provisionally allocated the Pasalar species to *Griphopithecus*³⁹. The third group comprises the European fossils that are mostly attributed to the genus *Dryopithecus*. The Dryopithecini are known from 12 to 8 Myr deposits in France, Austria, Germany, Spain and Hungary, and similar fossils from Georgia through to China have also been attributed to this genus.

The Afropithecini is characterized by thickening of the enamel on the molars (only actually measured on *Heliopithecus*³²), by broadening of the molars, reduction in molar cusps and ridges so that the occlusal surfaces become more flattened, hyperrobusticity of the canine, the breadth of which is almost as great as length, and enlargement of the premolars, particularly the uppers, which are 92–97% the size of the first molar. The premolar enlargement is greater than is seen in any other hominoid, living or fossil, except for the Moroto palate⁴⁰, which on this basis is also included in the Afropithecini^{32,36}. At least two premolar morphologies are included in this combination, *Afropithecus* and the Moroto palate having strong cusp heteromorphy on the third premolar and the Maboko '*Kenyapithecus*' having the cusps more nearly equal in size.

Enamel thickening of the molars and enlargement of the premolars seem likely to be linked to a dietary change in the afropithecine fossils. This may be an increased component of hard fruit in the diet, although the evidence for this is ambiguous⁴¹. The postcrania are only known for "*Kenya-*

pithecus" from Maboko Island, and they indicate little change from the generalized arboreal quadrupedalism present in the early Miocene hominoids like *Proconsul*⁴². Little is known of the ecological context of the afropithecine fossils, but while the best record is again from Maboko Island, there is some doubt about the ecological interpretation of this site. On the basis of the mammalian fauna, the hominoids appear to be associated with tropical forest, albeit a dry seasonal and probably deciduous forest⁴³. The land gastropods appear to indicate more arid conditions, although still with tree cover⁴⁴, indicating a high degree of climatic seasonality. It is hard to reconcile this interpretation with the apparent frugivorous adaptations of *Kenyapithecus* and the presence in the fauna of at least four other primates (one other hominoid, one monkey and two small-bodied apes^{45,46}), but the interpretation of seasonal forest is consistent with the little that is known about the other afropithecine sites as well.

The Kenyapithecini appear to have thicker enamel than the afropithecines^{47,48}, although enamel thickness has yet to be measured on any of the East African specimens. They occur in deposits slightly later than most afropithecines. Their postcranial remains are little different from those of the afropithecines (as determined by the Maboko postcrania⁴² for the latter). The kenyapithecines are represented by the humerus from Fort Ternan³⁶, undescribed remains from Pasalar and perhaps the Klein Hadersdorf specimens⁴⁹. The palaeoenvironments also appear to have been little different from those of the afropithecines, with closed woodland-forest indicated for Fort Ternan^{18,38,50}, although some authors infer open country or even grassland for this site⁵¹. At Pasalar, the evidence is strongly in favour of subtropical forest, probably evergreen but with a pronounced dry season⁵², and current work on the microwear of the Pasalar specimens, which are the oldest ones known to have thick enamel, indicates a diet of small hard objects, probably fruits⁵³.

The Dryopithecini retain the primitively thin enamel of early hominoids⁴⁷, and in some aspects of the skull and dentition it is little advanced over proconsulids. The subnasal morphology is slightly different from that of *Proconsul*, with some reduction in size of the incisive foramen, but there is no premaxillary lengthening similar to that of extant great apes. On the other hand, the morphology of the glabella region and the greater development of the brow ridges may indicate affinities with the

BOX 3 Relationship of *Sivapithecus* with humans or the orang-utan

Two options have been put forward in the literature: (1) *Sivapithecus* is related to humans⁵⁷, which has little support; and (2) *Sivapithecus* is related to the orang-utan⁵⁷, both of these may be combined with the third possibility that the orang-utan is the nearest living relative to humans⁵⁸, although there is much evidence against this proposition (Box 1). The characters proposed in support of these propositions are listed below.

Evidence for *Sivapithecus*–human relationship

1. Low cusped molars with thick enamel (although this is more likely to be an ancestral character for the pongine and hominine clades (Box 2))
2. robust mandibles with shallow, broad mandibular bodies;
3. reduced canines, upper canine being reduced mesiodistally, and with reduced canine sexual dimorphism;
4. tendency toward enlarged P3 metaconid.

Evidence for *Sivapithecus*–orang-utan relationship

1. Narrow interorbital distance;
2. broad high zygomatic region;
3. zygomatic foramen above lower rim of orbit;
4. smooth unstepped nasal floor, with elongated premaxilla, rotated antero-superiorly, with extremely narrow incisive canal;
5. facial profile concave with extreme airorhynch (Box 4);
6. great size discrepancy between the upper incisors.

A number of other characters shared by the orang-utan and *Sivapithecus* may also support this relationship, such as the lack of glabellar thickening and the shapes of the orbits and nose, but their significance is uncertain.

African apes and humans, but the polarity of these characters is uncertain³ (see Box 4). Postcranially, the limb bone shafts are rounded, the humerus shaft is straight with a convex deltoid region, and the elbow joint has increased ranges of extension and supination-pronation combined with stability, as seen in the distal humerus, which has a deep olecranon fossa, rounded capitulum and well developed medial and lateral ridges of the trochlea⁴⁹. These adaptations indicate upper arm and elbow function similar to that of the living great apes, with below-branch arboreal adaptations and elbow joints modified for stability⁴²; this might be evidence of a relationship, although the absence of some of these characters in *Sivapithecus*⁵⁴ may invalidate this conclusion.

The environment of the dryopithecines appears to have been forested. The vegetation suggested by floral evidence is of mesophytic forests, subtropical to warm temperate in nature, but probably partly deciduous and strongly seasonal⁵⁵. It is probable that the dryopithecine apes were mainly arboreal, but it is not clear at this stage what their diet was in such an environment. Their abundance in such sites as Rudabanya, Hungary, and Can Llobateres, Spain, and their wide distribution across the whole of southern Europe, demonstrates a successful adaptation spanning 3–4 Myr, but it is not possible with our present knowledge to say what was the basis of this success.

Ponginae

Over the past decade it has become accepted that some of the fossils formerly thought to be human ancestors are in fact on the line leading to the orang-utan. It has been shown earlier

that this line branched off before the separation of humans and the African apes, so that these fossils are now considered remote from human ancestry. '*Ramapithecus*' was once considered to be a human ancestor, but it is now viewed as a synonym of *Sivapithecus*⁵⁶. This has led some authors to make *Sivapithecus* a human ancestor⁵⁷, and others to recognize the affinities of these fossil genera with the orang-utan but provide evidence linking the orang clade with humans⁵⁸. The third option, and the one that I prefer⁵⁹, is that *Sivapithecus*, including '*Ramapithecus*', is related to the orang-utan and that they branch at node 2 in Box 1. The evidence for these options is summarized in Box 3.

The evidence supporting *Sivapithecus*–orang-utan relationship has been questioned recently on the basis of new postcranial remains of *Sivapithecus*. Two new humeri from Pakistan are said to retain⁵⁴ ancestral characters similar to those seen in *Proconsul* and *Kenyapithecus*, with the proximal shaft curving laterally and anteriorly (that is, convex laterally as in Old World monkeys), and a flat deltoid plane⁵⁴, whereas the distal articular surface has several derived great ape features. This combination of characters in *Sivapithecus* indicates that if it is indeed related to the orang-utan, characters like the straight shaft and convex deltoid plane that are present in the extant great apes must have arisen independently in the orang-utan and the African apes. Alternatively, it may be that *Sivapithecus* precedes the splitting of the great apes, so that the combination of cranial characters shared by it and the orang-utan must have arisen independently. In my view, the evidence of the face (Boxes 3 and 4) is of greater significance, for it involves several functional complexes of the

BOX 4 The characters of the homininae and the relationship of *Graecopithecus*

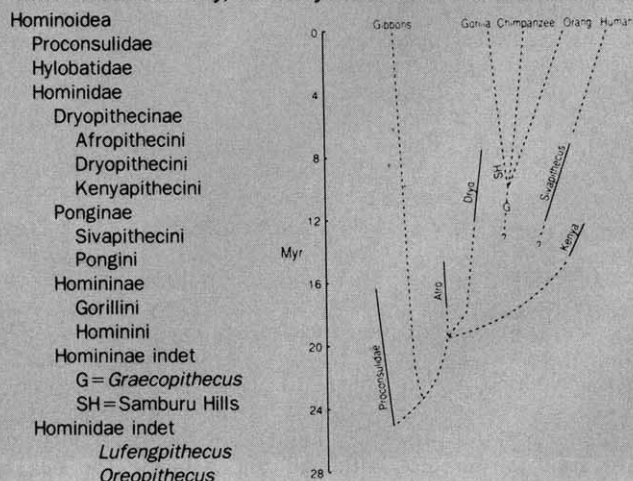
THE following characters have been claimed to distinguish the hominine branching point, node 3, Box 1, and their condition in fossil apes is described.

1. Prominent supra-orbital torus, continuous and bar-like in African apes and discontinuous in australopithecines: moderately prominent and discontinuous in *Dryopithecus*¹⁰⁵ and *Graecopithecus*⁶².
2. Prominent glabella: prominent in *Afropithecus*²⁶, *Dryopithecus*¹⁰⁵ and *Graecopithecus*⁶², and probably an ancestral hominoid character.
3. Frontal sinus present: present in *Proconsul*, *Afropithecus* and *Dryopithecus*, and so it is almost certainly an ancestral hominoid character. Not known if present in *Graecopithecus*, but may be so if it is associated with increased brow ridge development¹⁰⁶.
4. Elongated nasolabial clivus of the premaxilla with narrowing of the incisive foramen. Least developed in the gorilla but greater in chimpanzees and australopithecines, with formation of incisive canal⁷⁷. The hominine character state could either be gorilla-like lengthening of the premaxilla, which is not present in *Dryopithecus* but is in *Graecopithecus*, or it could be unlengthened as in *Dryopithecus*, but in both cases the elongated premaxilla and narrow incisive canal in *Graecopithecus* could constitute a hominine synapomorphy. The extreme reduction in diameter of the incisive canal and lengthening and rotation of the alveolar clivus of the premaxilla in the orang-utan are probably derived independently of the hominine condition.
5. Increased klinorhynch, which has been linked with most of the above characters in a single functional complex relating to the angle of hatching of the face to the cranium¹⁰⁶. It is difficult to identify this character in unsectioned skulls, but it is likely that some degree of airorhynch is primitive for hominoids, greatly exaggerated in the pongine lineage and modified towards klinorhynch in hominines, but as my observations lead me to believe that the bonobo is airorhynchous to some degree, it is difficult to be sure of the ancestral hominine character state. It is also highly doubtful that this character state can be identified in the broken and distorted partial skulls of *Dryopithecus* and *Graecopithecus*, although it is claimed that the former is klinorhynchous because of the brow ridge development¹⁰⁵. It is also claimed¹⁰⁷ that the molar cusp arrangement is related to cranial flexion, with the protocone being posteriorly placed relative to paracone in klinorhynchous skulls and transversely placed in airorhynchous forms, but this is the reverse of the usually accepted polarity for cusp orientation, with the primitive trigon having the protocone posteriorly placed.

6. Straight humeral shaft and convex deltoid plane: not known for *Graecopithecus* but present in *Dryopithecus*. Lateral and anterior curvature (that is, laterally and anteriorly convex) and flat deltoid plane represent the ancestral hominoid condition, as present in *Proconsul*, and these characters are present also in *Sivapithecus*⁵⁴. The degree of curvature, however, far exceeds that of other fossil hominoids, including *Proconsul*, and it may be a uniquely derived and functionally related characteristic of the sivapithecines.

7. Characters such as broader incisors and enlarged maxillary sinuses¹⁰⁵ are the only characters apart from the premaxillary/incisive complex (see character 4) that Begun¹⁰⁵ designates for the hominine clade, but I would interpret these as basal hominoid characters (node 2 on the cladogram in box 1), not diagnostic of the hominine clade. Three further characters¹⁰⁷ all appear to be ancestral, non-pongine characters not diagnostic of the hominine clade. These are: canine facets in line with molars (ancestral hominoid character), lingual cusps of the molars posteriorly placed (ancestral catarrhine character; see 5), and vertical orientation of incisors (non-pongine character).

Classification at family, subfamily and tribe level of the Hominoidea



nose, orbits and palate⁵⁹, and this supports the relationship of *Sivapithecus* with the orang-utan.

The presently accepted age range for *Sivapithecus* is 12.5 to 7 Myr⁶⁰, and it is this date that provides the earliest evidence for the splitting of the orang-utan from the other hominoids. Geographic distribution covers from IndoPakistan to Turkey. Palaeoenvironments were probably subtropical forests, strongly seasonal but most likely evergreen. On the evidence of postcrania, *Sivapithecus* was similar to other Miocene apes in being an arboreal quadruped⁵⁴ with perhaps a terrestrial component as well. Its diet appears to have been soft fruit⁴¹, which appeared to contradict the suggestion⁶¹ that thick enamel was adaptive for hard fruit diets; but the more recent conclusion that the microwear of the much earlier thick-enamelled hominoid from Pasalar⁵³ is indicative of a hard-fruit diet suggests that there was a dietary shift after the development of thick enamel⁴⁷, with dietary change independent of morphology.

Homininae

There are two groups of fossils that can be attributed to the Homininae, that is the African ape and human clade. These are the sample from Greece that has been called *Ouranopithecus macedoniensis* by some⁶², or synonymized with *Graecopithecus freybergi* by others⁶³, and the single maxilla from Samburu Hills²⁹ which has not yet been named. Both lack pongine characters, retaining the primitive condition for the characters that distinguish the pongine clade. The claim that on this basis the Greek fossils fit better with *Australopithecus afarensis*⁶² is certainly correct, but based as it is on primitive characters this conclusion has no phylogenetic significance⁶⁴. On the other hand, there are several characters linking *Graecopithecus* to the African ape and human clade⁶⁴. The supra-orbital torus is moderately well developed, with broad lateral orbital rims, but the torus is discontinuous, unlike the continuous bar-like structure in the African apes. The region of glabella is also well developed, although the condition in *Graecopithecus* is similar to that seen in dryopithecines and afropithecines, so that the significance of this character is uncertain⁶⁵. The premaxillary lengthening in *Graecopithecus* is a highly derived character shared with the African apes, and it is combined with the development of an incisive canal and an extended nasoalveolar clivus. These characters and some others are summarized in Box 4.

Hominidae indet

There are two groups of middle Miocene hominoid that have unknown relationships with the three subfamilies recognized here. These are *Lufengpithecus lufengensis* from the late Miocene of China and *Oreopithecus* from the late Miocene of Italy. The Chinese hominoid has been claimed to belong to the pongine clade⁶⁶, but the characters are of doubtful significance, for example the enlarged maxillary sinus, which may be an ancestral hominid character, or the size discrepancy between central and lateral incisors, which, although correct, is linked with an extremely low crowned and heavily buttressed morphology. Other supposed pongine characters of the zygomatic and orbital regions appear more likely to be primitive for the Hominidae.

Oreopithecus is one of the best known Miocene primates, and yet agreement has yet to be reached as to its taxonomic status. It is considered a separate family of the Cercopithecoidea by some⁶⁷ on the basis of its teeth, while others attribute it to a separate family of the Hominoidea⁶⁸ on the basis of its postcrania. All agree on the highly derived nature of the teeth, and it is the postcranial adaptations that are of particular interest, for *Oreopithecus* shares with living hominoids a number of apparent postcranial synapomorphies. Some of these characters are also present in *Dryopithecus*, and they have been used to support a relationship of this genus with the living great apes⁴⁹,

but their presence in *Oreopithecus*, which is so highly specialized cranially⁶⁸, must cast some doubt on this interpretation.

Hominoid evolutionary trends and the Homininae

The hominoid primates appear to have emerged during the Oligocene at least 25 Myr ago as a tropical African group inhabiting tropical forests in an equable climatic regime. The majority were arboreal frugivores occupying much the same niche as equivalent-sized monkeys today⁹. Two trends are apparent during the Miocene Period. The first led to increased enamel thickness on the molar teeth combined with little change in postcranial anatomy, and this is seen in the progression from *Proconsul* to *Afropithecus* to *Kenyapithecus*. This progression could be interpreted as an ancestor-descendent lineage, but the relationships of the later thick-enamelled hominoids to this putative lineage, *Graecopithecus* on the one hand and the pongine *Sivapithecus* on the other, are uncertain. These morphological changes may be interpreted functionally in dietary terms as a change to a diet with hard fruit as the main constituent, and this may be related in turn to the drier and more seasonal environments they lived in, but as there was little change in their postcranial skeletons it is likely that they retained their ancestral arboreal locomotor pattern.

The second trend produced change in the postcranial anatomy, leading to functional complexes similar to those of living great apes, but there was less change in the skull and particularly the teeth. There may have been some dietary change, despite the absence of change in enamel thickness, but this has yet to be tested. Postcranial change is seen in the progression from *Proconsul* and/or *Afropithecus* to *Dryopithecus* on the one hand, and to *Sivapithecus* on the other. *Dryopithecus* appears to share many of the great ape synapomorphies of the humerus, but these are present also in *Oreopithecus*, which makes their interpretation difficult. *Sivapithecus* combines some of these same characters of the distal humerus with primitive characters of the humeral shaft, and this confuses the issue still further if it is grouped in the orang-utan clade, for this indicates that some of the postcranial characters shared by the orang-utan and the African apes must have evolved independently. It is likely that locomotor and positional behaviours were related to environmental change, in this case the move from tropical forest habitats in Africa to the subtropical and warm temperate forest habitats of Europe, which would have been both more seasonal and more open, with less complex canopies.

The orang-utan lineage appears to have originated from within the first trend, with further modifications of skull and postcrania, but with little change in environments. There is no good evidence available at present to identify the origin of the African ape and human clade, the Homininae, but two alternatives can be mentioned. First, there may have been a continuation of the thick-enamelled trend independently of the pongine radiation⁴⁸, such as is seen in *Graecopithecus* for instance. The morphology of the alveolar clivus supports this alternative, in which case *Graecopithecus* can be interpreted as a fossil hominine, but there is a possibility that this morphology represents an ancestral great ape character which was primitively retained in *Graecopithecus* and the hominines (and was independently further modified to produce the more derived orang-utan condition).

The second alternative is that there may have been a continuation of the dryopithecine radiation. This is supported by the morphology of the postcrania, which is more similar in *Dryopithecus fontani* to the living great apes and humans than that seen in any other fossil hominoid⁴⁹. Two problems arise here, however, for *Oreopithecus* shares some of the same characters, and the pongine *Sivapithecus* is less advanced in this respect. *Oreopithecus* is clearly not a dryopithecine, which indicates these postcranial characters may be primitive for the great ape and human clade, but if *Sivapithecus* belongs in the orang-utan clade, as I have argued, the shared morphology of the orang-utan and the African apes must have arisen independently. This

weakens the phylogenetic potential for these postcranial characters in *Dryopithecus*, and although it could be suggested that *Dryopithecus* could belong in either of the two derivative clades (Ponginae or Homininae), additional supporting evidence would be needed before such a claim could be substantiated.

The earliest fossil evidence for the Homininae is based on fragments of jaw from latest Miocene to Pliocene deposits in Africa. An upper jaw from Samburu Hills²⁹ appears to have gorilla-like morphology of its teeth combined with thick enamel, which is very much what would be expected if thick enamel is identified as an ancestral hominine character. Its date is uncertain but is thought to have been about 9 Myr²⁹, and it is placed at this date on the phylogeny in Box 3 (identified as SH). Two lower jaws, one from Lothagam in northern Kenya is now dated at older than 5.6 Myr⁶⁹, and a slightly younger lower jaw from Tabarin⁷⁰ from deposits just over 5 Myr share very similar morphology with each other and with *Australopithecus afarensis*⁷¹.

It may be asked, however, to what extent this morphology would differ from that of thick-enamelled hominoid ancestors from the Miocene, and on what basis, therefore, should these fossils from Lothagam and Tabarin be identified as human ancestors. Early australopithecines are linked with living humans on the basis of shared characters related to bipedalism, but it has yet to be shown that their jaws and teeth differ from putative hominine ancestors. As the Tabarin and Lothagam jaws share only primitive characters with the earliest australopithecines, this is not enough to establish them definitely as human ancestors. The presence of some arboreal retentions in early australopithecines⁷², and their association with palaeoenvironments that may have been more wooded⁷³ than has formerly been thought, should also give rise to caution in identifying human origins⁷⁴. □

Peter Andrews is in the Department of Palaeontology, Natural History Museum, London SW7 5BD, UK.

1. Sarich, V. & Wilson, A. C. *Science* **158**, 1200-1203 (1967).
2. Pickford, M. & Andrews, P. *J. hum. Evol.* **10**, 11-33 (1981).
3. Andrews, P. in *Ancestors: the Hard Evidence* (ed. Delson, E.) 14-33 (Liss, New York, 1985).
4. Walker, A. C. & Pickford, M. in *New Interpretations of Ape and Human Ancestry* (eds Clöchon, R. L. & Corruccini, R. S.) 325-351 (Plenum, New York, 1983).
5. Walker, A. C. & Teaford, M. *Sci. Am.* **260**, 76-82 (1989).
6. Harrison, T. *J. hum. Evol.* **16**, 41-80 (1987).
7. Ward, C. V., Walker, A. C. & Teaford, M. *J. hum. Evol.* **21**, 215-220 (1991).
8. Walker, A. C., Falk, D., Smith, R. & Pickford, M. *Nature* **305**, 525-527 (1983).
9. Andrews, P. in *Aspects of Human Evolution* (ed. Stringer, C. B.) 25-61 (Taylor & Francis, London, 1981).
10. Ruff, C. B., Walker, A. C. & Teaford, M. *J. hum. Evol.* **18**, 515-536 (1989).
11. Walker, A. C., Teaford, M. F., Martin, L. & Andrews, P. *J. hum. Evol.* (in the press).
12. Andrews, P. *Bull. Br. Mus. nat. Hist. (Geol.)* **30**, 85-224 (1978).
13. Madden, C. T. *Primates* **21**, 241-252 (1980).
14. Boschetto, H. B., Brown, F. H. & McDougall, I. *J. hum. Evol.* **22**, 47-71 (1992).
15. Bishop, W. W., Miller, J. A. & Fitch, F. J. *Am. J. Sci.* **267**, 669-699 (1969).
16. Andrews, P. & Van Couvering, J. H. in *Approaches to Primate Paleobiology* (ed. Szalay, F. S.) 62-103 (Karger, Basel, 1975).
17. Collinson, M. *IAAP Newslett.* **8**, 4-12 (1985).
18. Andrews, P., Lord, J. & Evans, E. *Biol. J. Linn. Soc.* **11**, 177-205 (1979).
19. Aiello, L. in *Aspects of Human Evolution* (ed. Stringer, C. B.) 63-97 (Taylor & Francis, London, 1981).
20. Harrison, T. thesis, Univ. London (1982).
21. Kay, R. F. *Nature* **268**, 628-630 (1977).
22. Le Gros Clark, W. E. & Thomas, D. P. *Fossil Mammals Afr.* **3**, 1-27 (1951).
23. Rose, M. D. *J. hum. Evol.* **22**, 255-266 (1992).
24. Napier, J. R. & Davis, P. R. *Fossil Mammals Afr.* **16**, 1-69 (1959).
25. Wynn, T. & McGrew, W. C. *Man* **24**, 383-398 (1989).
26. Leakey, R. E. & Leakey, M. G. *Nature* **324**, 143-146 (1986).
27. Whybrow, P. J., McClure, H. A. & Elliott, G. F. *Bull. Br. Mus. nat. Hist. (Geol.)* **41**, 371-382 (1987).
28. Feibel, C. S. & Brown, F. H. *J. hum. Evol.* **21**, 221-225 (1991).
29. Ishida, H., Pickford, M., Nakaya, Y. & Nakano, Y. *Af. Study Monogr.* **1**, 73-86 (1984).
30. Delson, E. *Nature* **318**, 107-108 (1985).
31. Leakey, R. E. & Walker, A. C. *Nature* **318**, 173 (1985).
32. Andrews, P. & Martin, L. *Bull. Br. Mus. nat. Hist. (Geol.)* **41**, 383-393 (1987).
33. Conroy, G., Pickford, M., Senut, B., Van Couvering, J. & Mein, P. *Nature* **356**, 144-148 (1992).
34. Andrews, P. *Nature* **356**, 106 (1992).
35. Pickford, M. *J. hum. Evol.* **14**, 113-143 (1985).
36. Harrison, T. *Primates* (in the press).
37. Alpagut, B., Andrews, P. & Martin, L. *J. hum. Evol.* **19**, 397-422 (1990).
38. Pickford, M. *J. hum. Evol.* **16**, 305-309 (1987).
39. Martin, L. & Andrews, P. *Am. J. phys. Anthropol. suppl.* **12**, 126 (1991).
40. Pilbeam, D. *Bull. Peabody Mus. nat. Hist. New Haven* **31**, 1-185 (1969).
41. Teaford, M. & Walker, A. *Am. J. phys. Anthropol.* **64**, 191-200 (1984).
42. Rose, M. D. *J. hum. Evol.* **18**, 131-162 (1989).
43. Evans, E., Van Couvering, J. H. & Andrews, P. *J. hum. Evol.* **10**, 35-48 (1981).
44. Pickford, M. in *New Interpretations of Ape and Human Ancestry* (eds Clöchon, R. L. & Corruccini, R. S.) 421-439 (Plenum, New York, 1983).
45. Benefit, B. R. & McCrossin, M. L. *J. hum. Evol.* **18**, 493-497 (1989).
46. Harrison, T. *J. hum. Evol.* **18**, 537-557 (1989).
47. Andrews, P. & Martin, L. *Phil. Trans. R. Soc.* **334**, 199-209 (1991).
48. Martin, L. *Nature* **314**, 260-263 (1985).
49. Begun, D. *Am. J. phys. Anthropol.* **87**, 311-347 (1992).
50. Kappelman, J. *J. hum. Evol.* **20**, 95-129 (1991).
51. Retallack, G. *Am. J. phys. Anthropol.* **75**, 260 (1988).
52. Andrews, P. *J. hum. Evol.* **19**, 569-582 (1990).
53. King, T. thesis Univ. London (1992).
54. Pilbeam, D. R., Rose, M. D., Barry, J. C. & Shah, I. *Nature* **348**, 237-239 (1990).
55. Kretzoi, M., Krollop, E., Lorincz, H. & Palfalvy, I. *Földt. Int. Evi. Jel.* 365-394 (1974).
56. Kelley, J. & Pilbeam, D. in *Comparative Primate Biology* (ed. Swindler, D. R.) 361-411 (Liss, New York, 1986).
57. Kay, R. F. & Simons, E. L. in *New Interpretations of Ape and Human Ancestry* (eds Clöchon, R. L. & Corruccini, R. S.) 577-624 (Plenum, New York, 1983).
58. Schwartz, J. H. *Nature* **308**, 501-515 (1984).
59. Andrews, P. & Cronin, J. E. *Nature* **297**, 541-546 (1982).
60. Kappelman, J. et al. *J. hum. Evol.* **21**, 61-73 (1991).
61. Kay, R. F. *Am. J. phys. Anthropol.* **55**, 141-151 (1981).
62. Bonis, L. de, Bouvraïn, G., Geraads, D. & Koufos, G. *Nature* **345**, 712-714 (1990).
63. Martin, L. & Andrews, P. *Cour. Forsch. Senckenberg* **69**, 25-40 (1984).
64. Andrews, P. *Nature* **345**, 664-665 (1990).
65. Martin, L. in *Major Topics in Ape and Human Evolution* (eds Wood, B., Martin, L. & Andrews, P.) 161-187 (Cambridge University Press, UK, 1986).
66. Schwartz, J. *J. hum. Evol.* **19**, 591-605 (1990).
67. Szalay, F. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
68. Harrison, T. *J. hum. Evol.* **15**, 541-581 (1986).
69. Hill, A., Ward, S. & Brown, F. *J. hum. Evol.* **22**, 439-451 (1992).
70. Ward, S. & Hill, A. *Am. J. phys. Anthropol.* **72**, 21-37 (1987).
71. White, T. *Anthropos. Brno* **23**, 79-90 (1986).
72. Stern, J. & Susman, R. *Am. J. phys. Anthropol.* **60**, 279-317 (1983).
73. Andrews, P. *J. hum. Evol.* **18**, 173-181 (1990).
74. Wood, B. *Nature* **355**, 783-790 (1992).
75. Delson, E. & Andrews, P. in *Phylogeny of the Primates* (eds Luckett, W. P. & Szalay, F. S.) 405-446 (Plenum, New York, 1975).
76. Andrews, P. & Martin, L. *J. hum. Evol.* **16**, 101-118 (1987).
77. Ward, S. C. & Pilbeam, D. R. in *New Interpretations of Ape and Human Ancestry* (eds Clöchon, R. L. & Corruccini, R. S.) 211-238 (Plenum, New York, 1983).
78. Yunis, J. J. & Prakash, O. *Science* **215**, 1525-1530 (1982).
79. Goodman, M., Braunitzer, G., Stangl, A. & Shrank, B. *Nature* **303**, 546-548 (1983).
80. Slightom, J. L., Chang, L., Koop, B. & Goodman, M. *Molec. Biol. Evol.* **2**, 370-389 (1985).
81. Sibley, C. G. & Ahlquist, J. E. *J. molec. Evol.* **26**, 99-121 (1987).
82. Miyamoto, M. M., Slightom, J. L. & Goodman, M. *Science* **238**, 369-373 (1987).
83. Ruvolo, M. & Smith, T. F. *Molec. Biol. Evol.* **3**, 285-289 (1986).
84. Koop, B., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. *Nature* **319**, 234-238 (1986).
85. Hasegawa, M., Kishino, H. & Yano, T. in *Statistical Theory and Data Analysis II* (ed. Matusita, K.) 1-113 (Elsevier, Amsterdam, 1988).
86. Holmes, E. C., Pesole, G. & Saccone, C. *J. hum. Evol.* **18**, 775-794 (1989).
87. Sibley, C. G., Comstock, J. A. & Ahlquist, J. E. *J. molec. Evol.* **30**, 202-236 (1990).
88. Dene, H. T., Goodman, M. & Prychodko, W. in *Molecular Anthropology* (eds Goodman, M. & Tashian, R. E.) 171-196 (Plenum, New York, 1976).
89. Marks, J. *Cytogenet. Cell Genet.* **34**, 261-264 (1982).
90. Bianchi, N. O., Bianchi, M. S., Cleaver, J. E. & Wolff, S. J. *J. molec. Evol.* **22**, 323-333 (1985).
91. Dutrillaux, B. *J. Reprod. Fertil.* **28**, 105-111 (1980).
92. Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. *J. molec. Evol.* **18**, 225-239 (1982).
93. Hixon, J. E. & Brown, W. M. *Molec. Biol. Evol.* **3**, 1-18 (1986).
94. Templeton, A. R. *Evolution* **37**, 221-244 (1983).
95. Templeton, A. R. *Molec. Biol.* **4**, 315-319 (1987).
96. Krajewski, C. & Dickerman, A. W. *Syst. Zool.* **39**, 383-390 (1990).
97. Hasegawa, M. & Kishino, H. in *New Aspects of the Genetics of Molecular Evolution* (eds Kimura, M. & Takahata, N.) 303-317 (Springer, Berlin, 1991).
98. Djan, P. & Green, H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8447-8451 (1989).
99. Smouse, P. E. & Li, W.-H. *Evolution* **41**, 1162-1176 (1987).
100. Holmquist, R., Miyamoto, M. M. & Goodman, M. *Molec. Biol. Evol.* **5**, 201-216 (1988).
101. Saitou, N. *Am. J. Phys. Anthropol.* **84**, 75-85 (1991).
102. Ruvolo, M., Disotelli, T. R., Allard, M. W., Brown, W. M. & Honeycutt, R. L. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1570-1574 (1991).
103. Gebo, D. L. et al. *J. hum. Evol.* **17**, 393-401 (1988).
104. Larson, S. G. *Am. J. phys. Anthropol.* **76**, 449-462 (1988).
105. Begun, D. *Science* **257**, 1929-1933 (1992).
106. Shea, B. T. in *Orang-utan Biology* (ed. Schwartz, J.) 233-245 (Oxford University Press, Oxford, 1988).
107. Dean, D. & Delson, E. *Nature* **359**, 676-677 (1992).

ACKNOWLEDGEMENTS: I thank L. Aiello, E. Andrews, R. Bernor, D. Cameron, E. Delson, T. Harrison, F. Clark Howell, L. Martin, C. Stringer and M. Wolpoff for critically reading the manuscript.

Sensitivity of the Indian monsoon to forcing parameters and implications for its evolution

Warren L. Prell* & John E. Kutzbach†

* Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

† Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

General-circulation-model simulations used to estimate the sensitivity of the Indian monsoon to changes in orbital parameters, the orography of Tibet–Himalaya, atmospheric CO₂ concentration and the extent of glacial-age surface boundary conditions show that increased elevations and increased summer solar radiation are most effective in strengthening the monsoon. Strong monsoons (similar to today's) can be induced by strong solar forcing only when the elevation is at least half that of today. These conditions may have been attained in the late Miocene.

THE Indian Ocean summer monsoon dominates the seasonal wind and precipitation patterns and the character of land vegetation in southern Asia^{1,2}. The strong southwesterly monsoon winds over the Arabian Sea also induce dramatic changes in the seasonal upwelling and productivity of the Arabian Sea. The long-term evolution and variability of the monsoon can be attributed to changes in at least four types of large-scale forcing or boundary conditions: orbital parameters, mountain-plateau orography (tectonic), glacial-age surface boundary conditions and atmospheric CO₂ concentration. These amplify or dampen the seasonal processes of heating, latent heat transport and moisture convergence over the Asian continent, and thereby modify the strong summer monsoon circulation, which is characterized by high precipitation over southern Asia and strong southwesterly winds and upwelling in the Arabian Sea.

Here we use the results of general circulation models to show that the monsoon is most sensitive to elevation and radiation (orbital) changes. We then estimate their contributions to the long-term evolution and variability of the monsoon circulation, and conclude that high elevations (at least half modern) are a prerequisite for strong monsoons. Comparisons to palaeoclimate data indicate that these conditions were achieved in the late Miocene.

Sensitivity experiments with climate models

To identify the sensitivity of monsoon responses to various forcing and boundary condition factors, we synthesized the results of 16 experiments^{3–8}. These experiments used two versions of the NCAR Community Climate Model (CCM0 and CCM1) described in refs 9 and 10, respectively. The CCM0 experiments were for perpetual January and July conditions, and used prescribed sea surface temperature (SST), soil moisture and snow cover. The CCM1 experiments were for a full seasonal cycle and used an interactive mixed-layer ocean and interactive soil moisture and snow cover¹¹.

As a measure of the large-scale summer monsoon response relative to the respective control cases, we compare changes in the July values of surface temperature (ΔT , °C) as well as percentage changes of precipitation ($\% \Delta P$, mm d⁻¹), precipitation minus evaporation ($\% \Delta(P - E)$, mm d⁻¹) over southern Asia (Fig. 1) and the southerly component of the winds ($\% \Delta v$, m s⁻¹) in the western Arabian Sea (Fig. 1). We describe the sensitivity of the climate variable to changes of forcing by a relation of the form $\Delta y = b \Delta x$, where Δx is the fractional or percentage change in forcing factor x and Δy is the fractional or percentage change in climate variable y . The sensitivity coefficient, b , describes the magnitude (and direction) of the climatic response to forcing, provided that the response is roughly linearly related to the forcing.

Orbital forcing. Studies of a wide variety of monsoon indicators show that much of the 10³–10⁵-yr variability of the monsoon results from direct orbital forcing and associated internal climate feedbacks^{12–14}. To compare the modelled monsoon responses forced by different orbital parameters, we used the departure from modern values of the average Northern Hemisphere solar radiation for June, July and August (ΔS , measured as a percentage) (Fig. 2). This is a broad index of the seasonal insolation in the monsoon region rather than a specific index for a

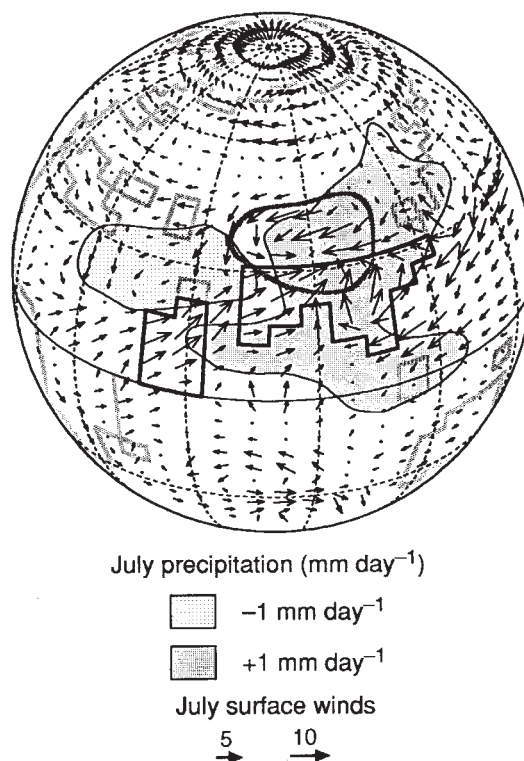


FIG. 1 Location of southern Asia and the western Arabian Sea area averages (bold line rectilinear enclosures). The smooth solid line over Tibet outlines the model elevations >1,500 m. Also shown are the changes (M - NM) for the winds (arrows, 5 and 10 m s⁻¹) and precipitation for July for the no-mountain CCM1 experiment. Areas of increased (>1 mm d⁻¹) and decreased (<1 mm d⁻¹) precipitation are shown by more dense and less dense shading, respectively. The western part of southern Asia (near Pakistan) becomes drier as the mountain elevations and monsoon precipitation increase to the east.

particular month and latitude. The monsoon response to changes in orbital parameters is estimated from an ensemble of CCMO experiments^{3,4,15} which use the orbital conditions appropriate for 3, 6, 9, 115 and 126 kyr ago (a range in Northern Hemisphere summer radiation of -5.8% to $+12.5\%$ relative to modern). We have also included a 6 kyr experiment that used the seasonal version of CCM1. This range of summer radiation incorporates most of the variation that has occurred because of orbital changes over the past 5 Myr (Fig. 3a). The 126 kyr experiment ($\Delta S = 12.5\%$) has $\sim 30\%$ higher precipitation (ΔP) over southern Asia than the control (Table 1, Fig. 2b), a strong monsoon response. Conversely, the 115 kyr experiment ($\Delta S = -5.8\%$) has a ΔP of $\sim 7\%$ less than the control, a weak monsoon response.

The monsoon precipitation responds roughly linearly to changes in solar radiation (Fig. 2b) with a slope (or sensitivity coefficient, b) of ~ 2.0 : a 1% increase in solar radiation produces roughly a 2% increase in ΔP . This sensitivity coefficient is slightly different from our previous estimate⁴ because we now use results of longer numerical experiments and a slightly different averaging area. The total range of ΔP is $\sim 40\%$ over this range of orbital changes (Fig. 2b). This range would be slightly higher during intervals of greater orbital forcing.

The range of temperature change (ΔT) over southern Asia is small ($< 2^\circ\text{C}$) and linear (Fig. 2a). The response of $\Delta(P-E)$ is also linear with a range of $\sim 40\%$ (Fig. 2c). The Arabian Sea winds (Δv) show a large response to orbital changes, ranging from $+35\%$ (6 kyr) to -40% (115 kyr) (Fig. 2d).

Glacial-age surface boundary condition forcing. The extent of glacial-age surface and atmospheric boundary conditions also affects the monsoon system⁴. During glacial periods, the Earth's surface differs from the present in several respects: sea surface temperature, sea level, albedo of the land surface, and the extent and elevation of large ice masses^{16,17}. In addition, atmospheric CO_2 levels are lower than in interglacials. Changes in sea surface temperature, land albedo and atmospheric carbon dioxide concentration are likely to be more important 'glacial' forcing factors

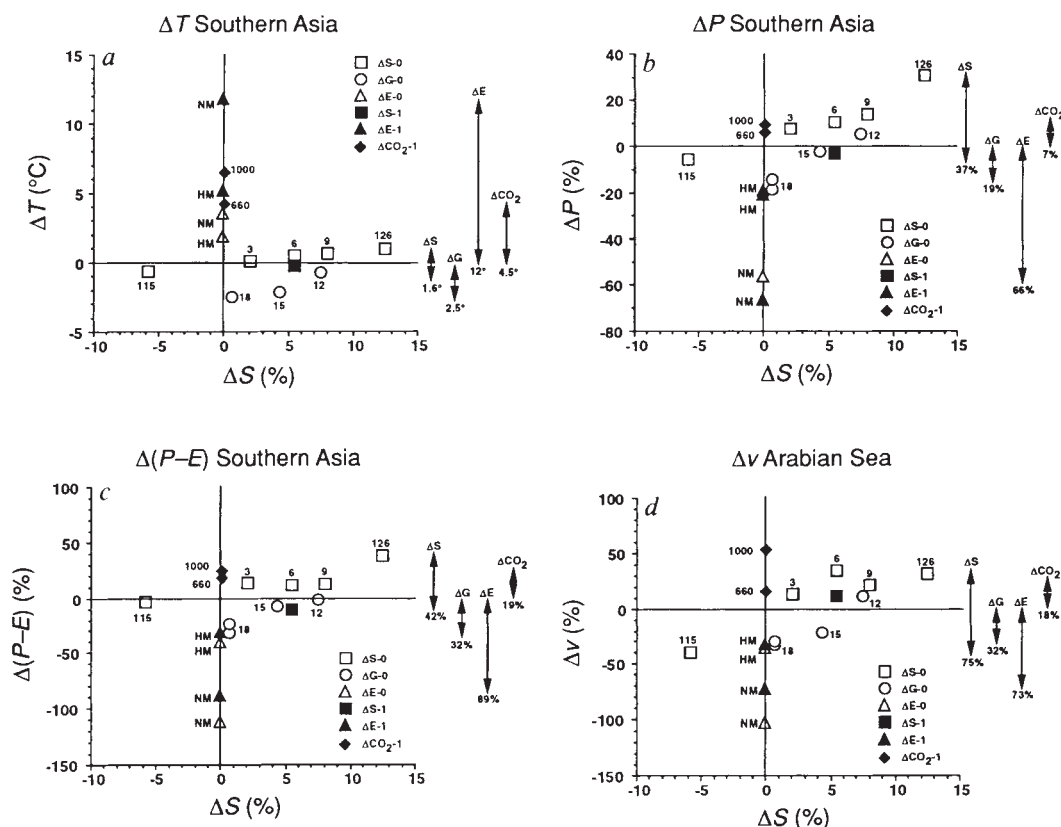
TABLE 1 Monsoon responses to different boundary conditions

	$\Delta S(\%)$	Southern Asia $\Delta T(^{\circ}\text{C})$	$\Delta P(\%)$	Arabian Sea $\Delta(P-E)(\%)$	$\Delta v(\%)$
Orbital radiation experiments					
Control-0	0.0	26.3 $^{\circ}\text{C}$	9.31 mm d^{-1}	4.44 mm d^{-1}	4.80 m s^{-1}
3 kyr	2.1	0.1	8	13	14
6 kyr	5.5	0.5	11	12	35
6 kyr-1	5.5	-0.3	-3	-10	12
9 kyr	8.0	0.6	14	13	21
115 kyr	-5.8	-0.7	-6	-4	-39
126 kyr	12.5	0.9	31	38	31
Glacial boundary condition experiments					
Control-0	0.0	26.3 $^{\circ}\text{C}$	9.31 mm d^{-1}	4.44 mm d^{-1}	4.80 m s^{-1}
12 kyr	7.5	-0.7	6	-1	12
15 kyr	4.3	-2.2	-2	-7	-22
18 kyr- CO_2	0.7	-2.5	-19	-32	-32
18 kyr-ice	0.7	-2.5	-15	-23	-30
Mountain-plateau experiments					
Control-0	0.0	26.3 $^{\circ}\text{C}$	9.17 mm d^{-1}	4.32 mm d^{-1}	4.61 m s^{-1}
Control-1	0.0	25.6 $^{\circ}\text{C}$	9.03 mm d^{-1}	5.18 mm d^{-1}	5.66 m s^{-1}
No mountain	0.0	3.5	-57	-112	-103
No mountain-1	0.0	11.8	-66	-89	-73
Half mountain	0.0	1.8	-19	-40	-36
Half mountain-1	0.0	5.1	-21	-31	-32
CO_2 experiments					
Control-1	0.0	30.9 $^{\circ}\text{C}$	5.69 mm d^{-1}	2.47 mm d^{-1}	4.05 m s^{-1}
CO_2 -660-1	0.0	4.3	7	19	18
CO_2 -1,000-1	0.0	6.6	10	26	54

Solar radiation changes (ΔS) and monsoon responses (experiment minus control) are given for July over southern Asia and the western Arabian Sea. Symbols and units for control case values are given in standard units. Experiment values are given as departures ($^{\circ}\text{C}$) from control for ΔT and percentage departures from the control for ΔP , $\Delta(P-E)$ and Δv . Experiments with -1 following the name were conducted with CCM1; all others were conducted with CCM0. 18 kyr-ice refers to a glacial experiment with a smaller ice sheet.

for the monsoon than details of the extent or height of the North American or Scandinavian ice sheets¹⁸. However, the extent of seasonal snow and ice over Tibet and Eurasia will affect the timing and strength of the monsoon¹⁹. The extent of these

FIG. 2 Sensitivity plots of changes in (a) surface temperature (ΔT , $^{\circ}\text{C}$); (b) precipitation (ΔP , %); (c) precipitation minus evaporation ($\Delta(P-E)$, %) for the southern Asian area (Fig. 1); and (d) changes in the southerly winds (Δv , %) over the western Arabian Sea (Fig. 1). All changes are shown relative to changes in solar radiation for the Northern Hemisphere summer (ΔS , %). Symbols indicate experimental boundary conditions and versions of the CCM. Open symbols, CCM0; solid symbols, CCM1. Squares, radiation experiments, ΔS (3 kyr, 6 kyr, 9 kyr, 115 kyr and 126 kyr). Circles, glacial boundary condition experiments, ΔG (18 kyr, 15 kyr, 12 kyr). Triangles, mountain elevation experiments, ΔE (NM, HM). Diamonds, CO_2 experiments, ΔCO_2 (660 p.p.m., 1,000 p.p.m.). Arrows at the right of each diagram summarize the range of the climate variable induced by each of the forcing factors. See Table 1 for summary of the experiments and the numerical results.



conditions has varied substantially over the past few million years and can be expected directly or indirectly to affect the monsoon system. Although observations indicate that the Indian Ocean monsoon was generally weaker during the Last Glacial Maximum^{12,13,20–23}, some glacial intervals do contain strong monsoons^{14,24}.

The monsoon response associated with full glacial boundary conditions (18 kyr) and deglacial conditions (15 kyr, 12 kyr) is estimated from experiments that used the CLIMAP^{16,17} boundary conditions for the last glacial maximum (sea surface temperature, land ice, sea ice, land albedo, atmospheric carbon dioxide concentration), but with appropriate modifications for deglacial conditions and for the changing solar radiation^{3,6}.

The influence of glacial conditions relative to the present is represented by the symbol ΔG where ΔG at 18 kyr is taken as 100%. Over southern Asia, the full glacial CCM0 experiment

(18 kyr) gives temperatures $\sim 2.5^\circ\text{C}$ lower (Fig. 2a) and precipitation $\sim 20\%$ lower with only small differences for experiments with lower ice sheets (Table 1, Fig. 2b). The deglacial experiments (15 kyr, 12 kyr) give smaller temperature changes (Fig. 2a), and precipitation changes are $\sim 10\text{--}15\%$ lower than orbital experiments (6 kyr and 9 kyr) with comparable radiation values (Fig. 2b). In other words, the experiments with glacial to deglacial boundary conditions yield similar estimates of glacial sensitivity coefficients after allowance is made for change in ΔS . At full glacial conditions, $\Delta(P - E)$ is depressed by $\sim 30\%$, but this departure decreases at higher insolation and reduced glacial conditions (Fig. 2c). Full glacial conditions also affect the southerly winds (Δv) which are reduced by 30%. The 15 kyr experiment also has weaker winds, and the 12 kyr winds are only $\sim 10\%$ lower than those in the experiments with similar ΔS (Fig. 2d). Although CCM0 may overestimate the change in glacial-age winds¹⁴, these changes are small relative to other forcing factors. These experiments indicate that the full range of glacial-age boundary conditions produces only about half the change of ΔP , $\Delta(P - E)$ and Δv than does the full range of orbitally-caused solar radiation change.

Mountain-plateau uplift forcing. The effects of tectonically induced orographic changes must also be understood if we are to unravel the monsoon's initiation and long-term evolution^{25,26}. Before the collision of India with Asia, the Himalayas and the Tibetan plateau did not exist in their present state, and the Asian continent was smaller. The smaller size and lower elevations of the pre-collision continent probably supported a lower heating contrast between land and sea because of less sensible heating and condensational heating over and on the flanks of the Tibetan plateau. Our CCM0 and CCM1 experiments with no mountains (NM), half-mountains (HM) and full-mountains (FM)^{5,27,28} explore the full range of elevation (ΔE) using modern mountain and plate location. The experiments do not address the effects of Andean-type mountains which may have existed on the southern margin of Asia because the spatial resolution of the model is not fine enough to represent narrow mountain ranges accurately.

The sensitivity of the climate to mountain and plateau elevations is illustrated by the higher surface temperatures over southern Asia accompanying lower topography. In the CCM1 experiments, the surface temperatures were $\sim 12^\circ\text{C}$ higher for the NM case and about 5°C higher for the HM case (Fig. 2a). Temperatures decrease towards the present as mountain and plateau elevation increases.

The monsoon response is indicated by the increased precipitation in all experiments as elevation is increased (Fig. 2b). The CCM1 experiments give larger NM responses than CCM0, but similar HM responses (Fig. 2b). The range of orographically induced ΔP over southern Asia is $\sim 66\%$ (CCM1) and is slightly larger than the range of orbitally induced ΔP changes ($\sim 40\%$). Both NM experiments give decreased precipitation in the mid-east (Fig. 1). The response of $\Delta(P - E)$ is similar in sign, with CCM0 giving slightly larger responses than CCM1 (Fig. 2c). Increasing elevation increases the precipitation over southern Asia but decreases the temperature; therefore, both effects tend to enhance the positive difference between precipitation and evaporation and hence increase the soil moisture and runoff.

For the NM experiments, the Arabian Sea summer winds are much weaker and the southerly component (Δv) is almost eliminated (Fig. 2d), implying that seasonal upwelling would not occur in the western Arabian Sea. The HM experiments have southwesterlies firmly established but $\sim 30\%$ weaker than the control experiments (Fig. 2d). The HM winds (Δv) are roughly equivalent to the effects of full glacial conditions (18 kyr) and the most negative radiation forcing (115 kyr, Fig. 2d).

Carbon dioxide. Observations suggest that the atmospheric concentration of CO_2 was higher during the Tertiary than today and may have been equivalent to CO_2 concentrations twice modern values at ~ 15 Myr (ref. 29). Higher CO_2 levels might

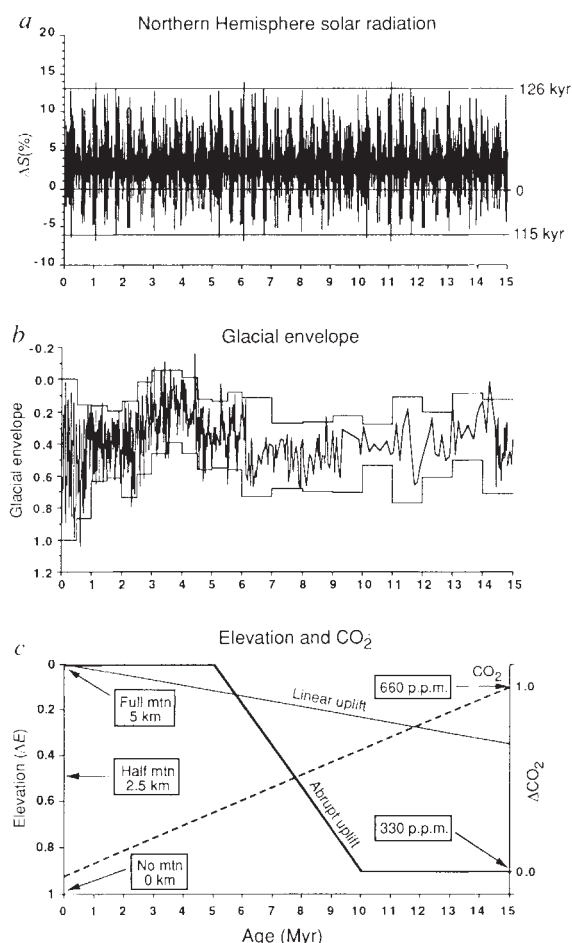


FIG. 3 Time series of the forcing factors for the sensitivity experiments. *a*, Changes (% difference from modern) in June–July–August solar radiation (ΔS) for the Northern Hemisphere, plotted at $\Delta t = 10$ kyr. Data are from Berger and Loutre³¹ and were calculated for the past 5 Myr and repeated to 15 Myr ago. *b*, The extent of glacial boundary conditions as estimated from the isotopic variation of tropical planktonic foraminifera³². The envelope of variability is scaled to full glacial conditions ($\Delta G = 1$ at 18 kyr ago, Last Glacial Maximum) and the modern control is scaled at 0. We use the upper envelope (minimum glacial boundary conditions) in our calculations to identify the initiation of strong monsoons. *c*, Variation of elevation and CO_2 over the past 15 Myr. CO_2 (dashed line) linearly decreases from double CO_2 (660 p.p.m.), with ΔCO_2 scaled at 1, to modern (330 p.p.m.) scaled at 0, over the past 15 Myr. The gradual uplift model and the abrupt uplift model (with rapid change during the late Miocene) are shown as paths of elevation against time (solid lines) over the past 15 Myr. Also indicated are the elevations in km represented by the no-mountain, half-mountain and full-mountain experiments with ΔE scaled at 1, 0.5 and 0, respectively.

be expected to strengthen the monsoon through a more active hydrological cycle. To examine the effect of increased CO₂ on the monsoon, we compared experiments with CO₂ levels of 330 p.p.m. (modern control), 660 p.p.m. (double CO₂) and 1,000 p.p.m. (triple CO₂)^{7,8}. The sensitivity of global temperature to the CO₂ level decreases with increasing CO₂, but the response is roughly linear for concentrations between 330 p.p.m. and 1,000 p.p.m. (ref. 7). For the double-CO₂ experiment, southern Asia is ~4 °C warmer, which is near the global average ΔT reported by Oglesby and Saltzman⁷. The monsoon indicators also increase slightly: ΔP by 7% and $\Delta(P-E)$ by 20%. The winds in the Arabian Sea are slightly stronger in the double-CO₂ experiment (Fig. 2d). In the triple-CO₂ experiment, temperatures are ~7 °C higher, ΔP is 11% higher, $\Delta(P-E)$ is 28% higher and the Arabian Sea winds are ~50% stronger.

The warming associated with CO₂ tripling is roughly equivalent to half the temperature effect associated with the NM experiments and more than twice the effect of the full range of glacial and orbital conditions. These warmer temperatures might allow some types of vegetation to survive at higher elevations than at present. Because these upper elevation limits would be lowered as CO₂ decreased, the amount of surface uplift that can be inferred from the elevation limits of some modern species (compared with their fossil equivalents) would be reduced. The effects of CO₂ increase on ΔP and $\Delta(P-E)$ are small, typically no more than half the orbitally forced changes of these variables. These experiments indicate that both temperature and precipitation decrease towards the present as the CO₂ concentration of the atmosphere decreases. The response to decreasing CO₂ is thus in contrast to the response to increasing mountain elevation, where temperature decreases but precipitation increases.

This sensitivity analysis shows that the monsoon response (ΔP) is enhanced by increases in solar radiation (ΔS), elevation (ΔE) and ΔCO_2 , but dampened by increases in glacial conditions (ΔG). If we assume (initially) that the effects of the four forcing factors are additive and independent, then the combined monsoon response (ΔP) for southern Asia can be estimated as the sum of four terms, $\Delta P = 2.0(\Delta S) - 20(\Delta G) - 66(\Delta E) + 10(\Delta\text{CO}_2)$, where ΔS is scaled in percentage departure from modern, and all other forcing factors are scaled from 1 to 0 as defined above (Figs 2, 3). To infer the history of monsoon variation, however, the time series of the forcing functions must also be estimated.

Estimates of temporal forcing

Just as the sensitivity of monsoon response to the four forcing factors is different (Fig. 2), so do the trends and variability of the four forcing factors differ over the past 15 Myr (Fig. 3). As the integrated effect of the four forcing factors is difficult to infer from theory or palaeoclimate data alone, we have combined the model sensitivity coefficients with known or postulated time series of the forcing factors to examine possible evolutionary paths of the monsoon system. Below, we present two reconstructions of the monsoon over the past 15 Myr, during which the modern monsoon system started up.

The orbitally induced radiation changes can be confidently estimated for the past 5 Myr (refs 30, 31). To approximate the statistical structure of orbital forcing, we constructed a time series of the percentage departure (from modern) of June–July–August solar radiation (ΔS) for the Northern Hemisphere for the past 5 Myr and then replicated the values twice to reach 15 Myr ago (Fig. 3a). This approach is taken because we are primarily interested in the form and periodicity of the orbital forcing rather than in the exact timing of individual cycles, which cannot be calculated with confidence beyond 5 Myr ago³¹. Three features of this time series are of particular interest: (1) the modern summer radiation is ~3% lower than the long-term average; (2) the 115 kyr and 126 kyr experiments capture more than 80% of the range of the maximum orbital variations over the past 5 Myr; and (3) the time series contains the characteristic

amplitude and periods of the orbital components (eccentricity, ~400 and 100 kyr; obliquity, ~41 kyr; precession, ~23 and ~19 kyr) that will be incorporated into the monsoon response.

The extent of glacial-age boundary conditions (and lowered sea level) for the past 15 Myr can be estimated from the oxygen isotope record of tropical planktonic foraminifera (Fig. 3b). We have used a reconstruction that normalizes values to the western tropical Pacific, an area assumed to have little or no temperature variation³². Unfortunately, the exact timing (relative to orbital variations) of glacially induced $\delta^{18}\text{O}$ variations is not well constrained in the Pliocene and late Miocene. Rather than estimating individual isotopic 'glacial' variations, therefore, we have estimated the envelope of glacial variability (Fig. 3b) by taking two standard deviations around the mean of 0.5-Myr or 1-Myr intervals (depending on data density). As we are interested in the temporal development of strong monsoons, we have used the upper limit of the envelope (minimum glacial conditions) for our time-series estimates of monsoon evolution. With that perspective, the magnitude of minimum glacial boundary conditions ranges between $\Delta G = 0$ (present) and $\Delta G \approx 0.3$ (7–10 Myr). Because the sensitivity of southern Asia ΔP to glacial conditions is relatively low, this term makes only a small contribution to the character of the monsoon history.

The selection of a temporal elevation curve for the Himalayas and Tibetan plateau is fraught with uncertainty^{25,33}, but evidence has emerged³⁴ for rapidly increased sedimentation, erosion, unroofing and climate change beginning near 10 Myr ago; see ref. 35. To illustrate the sensitivity of the monsoon evolution to elevation paths, we adopt two simple hypotheses: linear uplift, assuming constant gradual uplift from near sea level since 45 Myr ago, the time of initial collision (Fig. 3c); and abrupt uplift, assuming low elevations (<1 km or our NM experiment) from 15 to 10 Myr, followed by rapid increase to 5 km (full mountains) between 10 and 5 Myr (elevation reaches half the modern at about 8 Myr) and constant elevation since 5 Myr (Fig. 3c). Therefore, the normalized range of elevation (ΔE) since 15 Myr is close to 1.0 for the abrupt uplift case, but only about 0.3 for the gradual uplift case. We do not claim that these elevation hypotheses are the actual histories of elevation. However, the response of the monsoon to these elevation paths does provide some constraints on the elevations associated with evolution of the monsoon.

On the basis of the recorded changes in $\delta^{13}\text{C}$ of organic carbon²⁹, we adopt a linear decrease of atmospheric CO₂ concentration from double CO₂ (660 p.p.m.) at 15 Myr to modern values (330 p.p.m.) (Fig. 3c). For purposes of illustration, we normalize this range of ΔCO_2 to 1.0. According to the calculations in ref. 29, CO₂ concentrations may have been four times as high as present values 20 Myr ago so that the early portion of this time series (Fig. 3c) may include a period of rapid decrease from higher values.

This ensemble of estimated boundary conditions over the past 15 Myr, although not exact, gives a framework for evaluating the individual and combined importance of the four forcing factors on the development of the monsoon system.

Evolution of the monsoon

The estimated time series of monsoon precipitation (ΔP) over southern Asia reveals a highly variable monsoon with possible trends that would affect the interpretation of regional and global climate change (Fig. 4).

First, the estimates using gradual uplift (Fig. 4a) indicate monsoons that are stronger than the modern control throughout the entire 15-Myr interval, and almost all monsoons are considerably stronger than the Last Glacial Maximum (18 kyr) simulation. The occurrence of these strong monsoons is not supported by available palaeoclimate data.

Second, the estimates using abrupt uplift (Fig. 4b) give a large increase in monsoon strength (~60%) between 10 Myr and 5 Myr, with smaller variations since 5 Myr. Clearly, the choice

of elevation path determines when the monsoon is intensified. Because the orbital forcing is almost equal to the tectonic (elevation) forcing, cycles of strong and weak monsoons occur at all points in the elevation history, but strong monsoons in low elevation intervals are weaker than monsoons in the high elevation intervals.

Third, because the sensitivity of ΔP to ΔCO_2 is relatively small, the higher levels of CO_2 in the Miocene do not generate much stronger monsoons than at present in the abrupt uplift case (Fig. 4b), but do contribute somewhat in the gradual uplift case (Fig. 4a). The decrease of CO_2 toward the present actually weakens the monsoon response as both ΔT and ΔP decrease. This trend differs from the effect of increasing elevation which also decreases ΔT but considerably increases ΔP .

Fourth, because more extensive glacial conditions always act to dampen the monsoon, the global trend towards more glacial and cooler climates in the Tertiary produces weaker monsoon responses. However, the full glacial effect is only about half or less of the effects due to orbital or elevation changes. In general, the extent of glacial conditions has a significant effect only on short timescales when the coincidence of more extensive glacial conditions and decreased radiation would produce weaker monsoons. Because we used the upper envelope of glacial extent

rather than the full time-series variability, these weak monsoons are not illustrated in our time-series estimates of palaeo-monsoon strength.

The estimates of ΔP using abrupt uplift also indicate that monsoons as strong as the control ($\Delta P = 0$) occur consistently only after the elevation has risen to about half the modern elevation (Fig. 4b). Although monsoon variability is high because of the orbital forcing, higher elevations are essential for circulation as strong or stronger than the modern monsoon. In the abrupt uplift case, the mountains reach half their present elevation between 7 and 8 Myr. Several continental and oceanic indicators of monsoonal climates also indicate that the monsoon either started or intensified about this time. Marked shifts in the $\delta^{13}\text{C}$ and mammal populations in Pakistan at about 7–7.4 Myr, and enrichment of $\delta^{18}\text{O}$ soil carbonates and decreased thickness of leaching zones near 8 Myr, have been interpreted as indicating a large change in regional climate, accompanied by a shift to more seasonal, water-stressed C4 vegetation (Fig. 4d, from Quade *et al.*³⁶). The $\delta^{13}\text{C}$ shift is observed on other continents and is interpreted as reflecting the expansion of C4 vegetation due to decreasing atmospheric CO_2 (T. E. Cerling, personal communication). However, the $\delta^{18}\text{O}$ and leaching zones indicate a change in the regional climate of this monsoonal area. The global nature of the $\delta^{13}\text{C}$ shift invites the speculation that the increased elevations and monsoonal activity lead to increased weathering, which is one way to draw down the global CO_2 (ref. 37). Another possibility is that elevation-forced changes in wind patterns could have led to changes in ocean circulation and ocean carbon cycling. In the Arabian Sea, the appearance and abundance of planktonic foraminifera and radiolaria, which are endemic upwelling species, occurs at ~8 Myr (Fig. 4c)^{38–40}. Although the continental and marine indicators are not exactly synchronous, they suggest that the period around 8 Myr is a time of monsoon intensification and monsoon wind-induced upwelling. With our model-based estimate of climate sensitivity, therefore, the hypothesis of abrupt uplift in the late Miocene produces a rapid climate response that is consistent with the palaeoclimate observations. The hypothesis of gradual uplift does not fit these observations. We conclude that the elevation threshold required to induce a strong monsoon response was probably not exceeded until the late Miocene, and that this provides a potential constraint on the 'true' surface uplift and elevation history of the Tibetan plateau and Himalayas. □

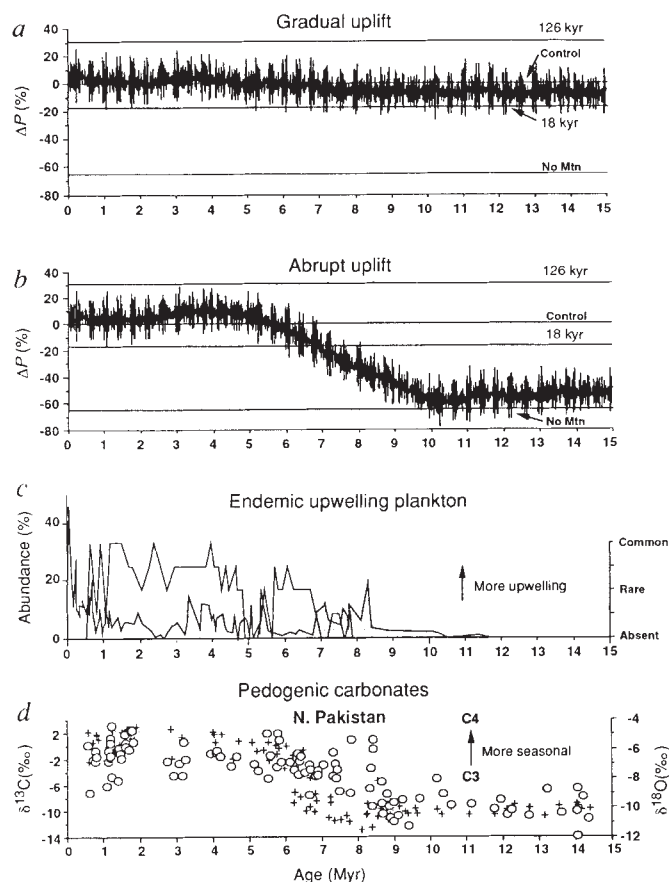


FIG. 4 Estimated time histories of changes in monsoonal precipitation over Southern Asia and observed monsoon indicators. *a*, ΔP history according to the gradual uplift model. Also shown are the ΔP estimates for key experiments (126 kyr, 18 kyr, NM and control, $\Delta P = 0$) to illustrate the range of model results compared to the total changes in estimated monsoon precipitation. *b*, ΔP history using the abrupt uplift model. *c*, Time series of endemic upwelling species of plankton: %foraminifera, *Globigerina bulloides* (shaded), and qualitative abundance of radiolarian *Actinomma* spp. as indices of upwelling induced by monsoon winds^{38,40}. *d*, Time series of carbon (crosses) and oxygen (circles) isotopes in soil carbonates in Pakistan as an index of a shift in regional climate and to C4 vegetation which is interpreted to indicate a more arid, highly seasonal environment³⁶.

Received 8 June; accepted 26 November 1992.

- Hastenrath, S. *Climate and the Circulation of the Tropics* (Reidel, Boston, 1985).
- Fein, J. S. & Stephens, P. L. (eds) *Monsoons* (Wiley, New York, 1987).
- Kutzbach, J. E. & Guetter, P. J. *J. Atmos. Sci.* **43**, 1726–1759 (1986).
- Prell, W. L. & Kutzbach, J. E. *J. geophys. Res.* **82**, 8411–8425 (1987).
- Kutzbach, J. E., Guetter, P. J., Ruddiman, W. F. & Prell, W. L. *J. geophys. Res.* **94**, 18393–18407 (1989).
- Kutzbach, J. E., Guetter, P. J., Behling, P. J. & Selin, R. in *Global Climates for 9000 and 6000 Years Ago in the Perspective of Glacial/Interglacial Climatic Change* (eds Wright, H. E., Kutzbach, J. E., Ruddiman, W. F., Street-Perrott, F. A. & Webb, T. III) (Univ. of Minnesota Press, Minneapolis, in the press).
- Oglesby, R. J. & Saltzman, B. *Geophys. Res. Lett.* **17**, 1089–1092 (1990).
- Oglesby, R. J. & Saltzman, B. *J. Clim.* **5**, 66–92 (1992).
- Pitcher, E. J. *et al. J. Atmos. Sci.* **40**, 580–604 (1983).
- Randel, W. J. & Williamson, D. L. *J. Clim.* **3**, 608–633 (1990).
- Covey, C. & Thompson, S. L. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **75**, 331–341 (1989).
- Prell, W. L. in *Climatic Processes and Climate Sensitivity*, *Geophys. Monogr. Ser.* 29 (eds Hansen, J. E. & Takahashi, T.) 48–57 (Am. geophys. Un., Washington DC, 1984).
- Prell, W. L. in *Milankovitch and Climate, Part 1* (eds Berger, A., Imbrie, J., Hays, J., Kukla, G. & Saltzman, B.) 349–366 (Reidel, Hingham, Massachusetts, 1984).
- Clemens, S., Prell, W., Murray, D., Shimmield, G. & Weedon, G. *Nature* **353**, 720–725 (1991).
- COHMAP Members *Science* **241**, 1043–1052 (1988).
- CLIMAP Project Members *Science* **191**, 1131–1137 (1976).
- CLIMAP Project Members *Geol. Soc. Am. Map and Chart Series No.* **36**, (1981).
- Manabe, S. & Broccoli, A. J. *J. geophys. Res.* **90**, 2167–2190 (1985).
- Barnett, T. P., Dumenil, L., Schlese, U., Roeckner, E. & Latif, M. *J. Atmos. Sci.* **46**, 661–685 (1989).
- Street-Perrott, F. A. & Harrison, S. P. in *Climatic Processes and Climate Sensitivity*, *Geophys. Monogr. Ser.* 29 (eds Hansen, J. E. & Takahashi, T.) 118–129 (Am. geophys. Un., Washington DC, 1984).
- van Campo, E., Duplessy, J. C. & Rossignol-Strick, M. *Nature* **296**, 56–59 (1982).
- Prell, W. L. & van Campo, E. *Nature* **323**, 526–528 (1986).
- Anderson, D. M. thesis, Brown Univ. (1990).
- Clemens, S. C. & Prell, W. L. *Paleoceanography* **5**, 109–145 (1990).
- Ruddiman, W. F., Prell, W. L. & Raymo, M. E. *J. geophys. Res.* **94**, 18379–18391 (1989).

26. Manabe, S. & Broccoli, A. J. *Science* **247**, 192-195 (1990).
 27. Ruddiman, W. F. & Kutzbach, J. E. *J. geophys. Res.* **94**, 18409-18427 (1989).
 28. Kutzbach, J. E., Prell, W. L. & Ruddiman, W. F. *J. Geol.* (in the press).
 29. Arthur, M. A. in *Advisory Panel Report on Earth System History* 51-74 (NSF, Division of Ocean Studies, 1991).
 30. Berger, A. L. *Quat. Res.* **9**, 139-167 (1978).
 31. Berger, A. & Loutre, M. F. *Quat. Sci. Rev.* **10**, 297-317 (1991).
 32. Prentice, M. L. & Matthews, R. K. *J. geophys. Res.* **96**, 6811-6827 (1991).
 33. Molnar, P. & England, P. *Nature* **346**, 29-34 (1990).
 34. Raymo, M. & Molnar, P. in *AGU 1991 Fall Mtg Prog. Abstr.*, *Eos* **257** Suppl. (Am. Geophys. Un., Washington DC, 1991).
 35. Harrison, T. M., Copeland, P., Kidd, W. S. F. & An Yin, *Science* **255**, 1663-1670 (1992).
 36. Quade, J., Cerling, T. E. & Bowman, J. R. *Nature* **342**, 163-166 (1989).
 37. Raymo, M. E. & Ruddiman, W. F. *Nature* **359**, 117-122 (1992).
 38. Prell, W. L., Murray, D. W., Clemens, S. C. & Anderson, D. M. in *The Indian Ocean: A Synthesis of*

- Results from the Ocean Drilling Program Geophys. Monogr.* 70 (eds Duncan, R. A., Rea, D. K., Kidd, R. B., von Rad, V. & Weissel, J. K.) (Am. geophys. Un., Washington DC, in the press).
 39. Kroon, D., Steens, T. & Troelstra, S. R. in *Proc. ODP Scient. Res.* **117** (eds Prell, W. L. et al.) 257-263 (Ocean Drilling Project, College Station, Texas, 1991).
 40. Nigrini, C. in *Proc. ODP Scient. Res.* **117** (eds Prell, W. L. et al.) 89-126 (Ocean Drilling Program, College Station, Texas, 1991).

ACKNOWLEDGEMENTS. We thank R. Oglesby for the results of his CO₂ experiments, P. Guetter for running the other model experiments, P. Behling, R. Selin and P. Howell for summaries and graphics of simulations and observations, D. Murray and A. Martin for Arabian Sea data, and J. Donnelly for critical reading of the manuscript. This work was supported by grants from the NSF Ocean Science Program to Brown University and to the University of Wisconsin-Madison by the NSF Climate Dynamics Program. Sensitivity analyses were also supported by a DOE grant to Brown University. The computations at the National Center for Atmospheric Research (NCAR) were sponsored by the NCAR Computing Facility.

LETTERS TO NATURE

Magnetic fields as an alternative explanation for the rotation curves of spiral galaxies

E. Battaner*, J. L. Garrido*, M. Membrado†
& E. Florido*

* Departamento de Física Teórica y del Cosmos,
University of Granada, Spain

† Departamento de Física Teórica, University of Zaragoza, Spain

THE flat rotation curves of spiral galaxies are usually regarded as the most convincing evidence for dark matter. The assumption that gravity alone is responsible for the motion of gas beyond the visible disks of galaxies leads directly to the conclusion that there must be perhaps 10 times as much dark matter as visible matter. Other forces besides gravity are usually neglected, as order-of-magnitude arguments seem to suggest they cannot be important. The existence of dark matter is, however, so important an issue that we believe it is wise to consider other possibilities. Here we argue that an azimuthal magnetic field can carry slightly ionized gas with the general galactic rotation, rendering dark matter unnecessary (a related idea was first proposed by Nelson¹). For the illustrative case of M31, a magnetic field of 6 μ G is required, and the synchrotron emission of relativistic electrons in this field is compatible with the observations.

The outermost regions of disc galaxies are observed by means of their gas content; in particular the rotation curve is obtained from the Doppler shift of the 21-cm line of interstellar atomic hydrogen. This gas is sufficiently ionized for magnetic field lines to be frozen in. Magnetic forces therefore influence the motion and distribution of the observed atomic hydrogen. Here we will investigate whether this influence is large enough. The basic assumption behind our equations is that magnetic tension holds the gas together, forcing it to spin faster than pure gravitation would.

Magnetic fields noticeably affect the behaviour of the peripheries of galaxies. Intergalactic magnetic fields may produce a warping of the disc^{2,3}, corrugations⁴ or enhanced ellipticity of the outermost isophotes⁵. Interstellar magnetic fields may have a considerable influence on galactic evolution⁶, and in particular may be essential to maintain external rings^{7,8}. Here we follow refs 7 and 8 in introducing the equations that will be used in our calculations.

The large-scale structure of the magnetic field may be either 'circular' (pure azimuthal field), or a structure with field lines open to space⁹. In either case the azimuthal component of the magnetic field is large, so only the effect of this component is considered here. We adopt the equations proposed in ref. 7 with one important modification. It was considered there that the stellar and the gaseous systems corotate. This condition must now be removed. For stars in the outermost disc the gravitational

and centrifugal forces are in balance. This cannot be assumed for the gas: in the absence of dark matter, the gravitational forces are due to stars, but gas and stars do not corotate. In the outermost disc, stars are not observed, and we may assume that the stellar rotation curve is keplerian. The rotation curve of the gas, however, remains flat, and we do not assume keplerian rotation.

The radial component of the motion equation simply becomes

$$\rho \left[-\frac{\partial V}{\partial r} + \frac{v^2}{r} \right] - \frac{\partial p}{\partial r} - \frac{1}{8\pi r^2} \frac{\partial(r^2 B^2)}{\partial r} = 0$$

where V is the gravitational potential, v the rotation velocity, p the gas pressure, B the azimuthal magnetic field, ρ the density and r the galactocentric radius.

This equation may be integrated to yield

$$B^2 = B_g^2 + B_p^2 + B_0^2 \frac{r_0^2}{r^2}$$

where

$$\frac{B_g^2}{8\pi} = \frac{1}{r^2} \int_{r_0}^r \rho \left(v^2 \xi - \xi^2 \frac{\partial V}{\partial \xi} \right) d\xi$$

is the contribution of gravity and centrifugal forces to the required magnetic field,

$$\frac{B_p^2}{8\pi} = -\frac{1}{r^2} \int_{r_0}^r \xi^2 \frac{\partial p}{\partial \xi} d\xi$$

is the contribution of the pressure gradient force, and the last term represents the influence of the boundary conditions. B_0 is the magnetic strength at the boundary radius r_0 , which was taken as 5 kpc. The last term is proportional to r^{-2} , so the results obtained for large radii do not depend very much on the choice

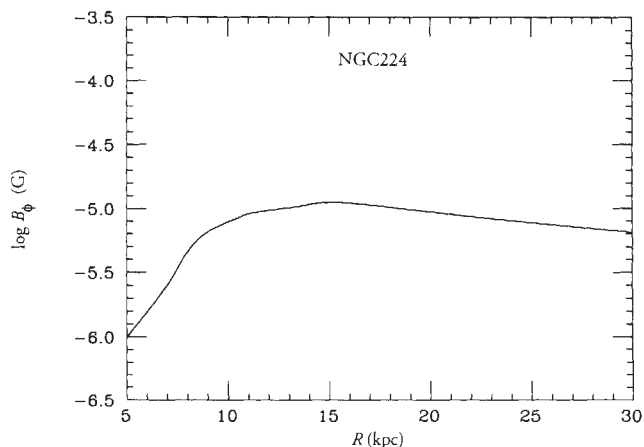


FIG. 1 Azimuthal component of the magnetic field strength in M31.

of B_0 and r_0 . Preliminary calculations confirmed this. For $r < 20$ kpc, the adopted densities were the same as in refs 7 and 8, from atomic hydrogen and carbon monoxide data. The densities for $r > 20$ kpc were obtained assuming an exponential disc where the density decreases by one order of magnitude in 10 kpc. The gravitational potential was taken from ref. 10, based on the luminosity profile. The rotation velocities were taken from refs 10 and 11.

The results for magnetic field strength are shown in Fig. 1. For B_0 , a value of 10^{-6} G was assumed, as the higher observational values are found in a ring centred at ~ 10 kpc (ref. 12). This ring was also found in our calculations, although the absolute values were higher and the profile was shifted outwards. The value and slope at 30 kpc are more important for our present purposes. We found $B_0 = 6 \times 10^{-6}$ G and $dB/dr = -2 \times 10^{-7}$ G kpc $^{-1}$.

For easier comparison with observations¹², we also calculated the synchrotron radiation produced by these magnetic fields and relativistic electrons, following the method of ref. 8. We assumed that the number of relativistic electrons per unit volume and energy interval is proportional to ρ/B , to take into account the energy loss of cosmic electrons due to the synchrotron emission itself (this is based on the model of ref. 13). The constant of proportionality was adopted to match the brightness temperature at 10 kpc. The theoretical profile (Fig. 2) does not agree well with observations for radii less than 10 kpc, but values at small radii depend on boundary conditions, and vertical magnetic fields probably become important. We focus instead on the peripheral results, where the agreement is good. This shows that our values for the magnetic field in the outermost disc are compatible with observations.

To interpret the results better, it is possible to integrate the above equation analytically for large radii, with the following assumptions: the gravitational potential is that created by a point-like galactic mass; the profile of gas rotational velocity is flat; the density and the pressure decrease exponentially. A function $B(r)$ is then easily found. Asymptotically, for radii several times the scale length of the exponential disc, this function is approximated by $B = \text{constant} \times r^{-1}$. This is essentially the B -profile that renders the outer rotation curve flat. The function $B(r)$ plotted in Fig. 1 is fitted by this simple law at large radii.

The obtained magnetic strength is reasonable for the magnetic ring, but could be considered too large at 30 kpc. The stability of the disc with such a strong field and the stability of the field itself pose some problems. Turbulent diffusion and the development of Parker's instabilities in a disc with high magnetic pressure would tend to dissipate the field rapidly. Is some dynamo mechanism enhanced at the rim to compensate these losses?

Are 'corrugations', mainly observed in the outer regions of discs¹⁴, a result of enhanced production of Parker's instabilities? In any case, the required magnetic fields are compatible with observations of the radio continuum. Also, these fields must connect with those in the intergalactic medium, which may reach values higher than 10^{-6} G (ref. 15).

We conclude that the flatness of the rotation curve for the outer disc of spiral galaxies, generally presumed to be evidence of dark matter, could instead be the result of interstellar magnetic fields. Direct measurements of these fields by means of the polarized synchrotron radio continuum, at large galactocentric distances, are needed to confirm this interpretation. Magnetic field strengths of $\sim 6 \times 10^{-6}$ G combined with a slow radial decrease could produce a non-keplerian slope for the rotation velocity. $\square$

Received 10 August; accepted 3 November 1992.

1. Nelson, A. H. *Mon. Not. R. astr. Soc.* **233**, 115–121 (1988).
2. Battaner, E., Garrido, J. L., Sanchez-Saavedra, M. L. & Florido, E. *Astr. Astrophys.* **251**, 402–410 (1991).
3. Binney, J. *Dynamics of Disc Galaxies in Warps* (ed. Sundelius, B.) 297–309 (Göteborgs Univ., Sweden, 1991).
4. Nelson, A. H. & Matsuda, T. *Mon. Not. R. astr. Soc.* **191**, 221–235 (1980).
5. Battaner, E., Florido, E. & Sanchez-Saavedra, M. L. *Astr. Astrophys.* **253**, 89–91 (1992).
6. Beck, R. *Interstellar Magnetic Fields in Interstellar Magnetic Fields: Past, Present and Future* (eds Beck, R. & Grave, R.) 2–7 (Springer, Berlin, 1986).
7. Battaner, E. & Sanchez-Saavedra, M. L. *Astrophys. J.* **302**, 450–458 (1986).
8. Battaner, E., Florido, E. & Sanchez-Saavedra, M. L. *Astrophys. J.* **331**, 116–123 (1988).
9. Beck, R. *Magnetic Fields and Extragalactic Objects in Interstellar Magnetic Fields in Normal Galaxies* (eds Asseo, E. & Gressillon, D.) 113–128 (Edition de Physique, Cargèse, 1987).
10. Kent, S. M. *Astron. J.* **93**, 816–832 (1986).
11. Newton, K. & Emerson, D. T. *Mon. Not. R. astr. Soc.* **181**, 573–590 (1977).
12. Beck, R. *Astr. Astrophys.* **106**, 121–132 (1982).
13. Lerche, I. & Schlickeiser, R. *Astr. Astrophys.* **116**, 10–26 (1982).
14. Florido, E., Battaner, E., Prieto, M., Mediavilla, E. & Sanchez-Saavedra, M. L. *Mon. Not. R. astr. Soc.* **251**, 193–198 (1991).
15. Vallee, J. P. *Astron. J.* **99**, 459–462 (1990).

ACKNOWLEDGEMENTS. We thank F. Cornet for suggesting the possibility explored in this paper. This work was in part supported by the Spanish DGICYT.

Long-term solar brightness changes estimated from a survey of Sun-like stars

G. W. Lockwood*, Brian A. Skiff*, Sallie L. Baliunas† & Richard R. Radick‡

* Lowell Observatory, 1400 W. Mars Hill Road, Flagstaff, Arizona 86001, USA

† Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, Massachusetts 02138, USA

‡ USAF Phillips Laboratory, Solar Research Branch, Sunspot, New Mexico 88349, USA

THE brightness of the Sun varies during the 11-year solar cycle, typically by less than 0.1% (refs 1, 2), and a larger brightness variation is thought to have occurred during the Maunder minimum, from AD 1645 to 1715 (refs 3–5). But because individual solar cycles are different in form, amplitude and length, and because accurate solar data have been available only for the most recent two or three cycles, there is no direct way of understanding long-term solar variability. Here we present a compilation of eight years of observations of 33 Sun-like stars and report year-to-year brightness changes that substantially exceed the analogous solar fluctuations. We have also measured chromospheric magnetic activity in these stars and find that it correlates with the brightness variations. During 1980–88, solar chromospheric variability was comparable to that observed in the stellar survey, but solar brightness variations were only one-quarter as large. This suggests that the Sun is in an unusually steady phase compared to similar stars, which means that reconstructing the past historical brightness record, for example from sunspot records, may be more risky than has been generally thought.

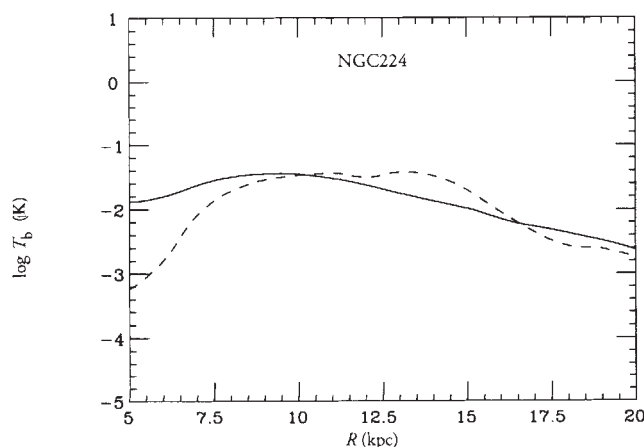


FIG. 2 Brightness temperature profiles for M31. Solid line corresponds to the observational results and dashed lines to this work.

Radiometry obtained since 1978 by the Nimbus 7 and Solar Maximum Mission satellites, spanning more than one 11-year solar activity cycle, showed that the solar total irradiance (the 'solar constant') decreased by 0.08% between the 1980 activity maximum and the 1986 minimum, and then began slowly to recover^{1,2}. These measurements permitted the first quantification of correlations between total irradiance and other solar activity indicators such as sunspot number, and these correlations in turn formed the basis for several reconstructions of solar irradiance values back as far as the Maunder Minimum, 1645–1715 (refs 3–5). But because individual solar activity cycles have very different forms, amplitudes and lengths, it is difficult to assess the validity of reconstructions depending on observations made during a single solar cycle.

Here we approach this question by comparing solar variability to its stellar analogues. Our stellar data comprise measurements of photometric brightness and spectrophotometric chromospheric emission for stars similar to the Sun but sampling ranges of surface temperature and mean chromospheric activity centred on present solar values. Many of these stars show Sun-like magnetic activity cycles. The 8-year interval of our observations is long enough to detect cycle-like variations in many of the stars, even though it is too short to cover the full cycle in every case.

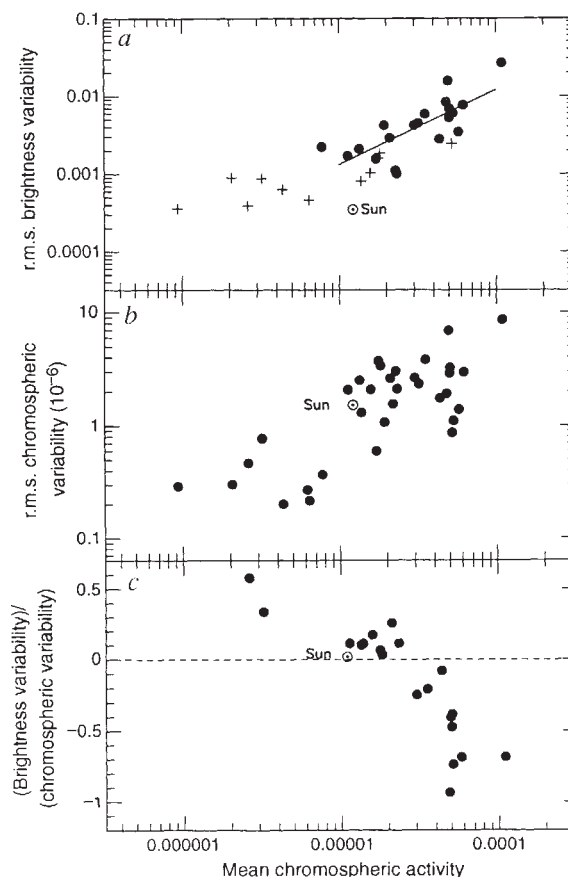
High-precision photometric observations, begun in 1984 using the 0.5-m telescope and associated equipment at the Lowell Observatory, serve as our substitute for solar irradiance data. These measurements were made through interference filters in 15–20-nm passbands at 472 nm ('b', blue) and 551 nm ('y', yellow) in the visual continuum for 33 programme stars and their nearby reference stars. Our use of redundant reference stars usually allowed us to decide unambiguously which stars were variable and by how much^{6,7}. Assuming that stars radiate as black bodies, observations made in the visual continuum (near the peak of the solar spectrum) exceed variations in total

(bolometric) irradiance by 20–25% for a star like the Sun, and we have adjusted the solar irradiance data accordingly. Because ultraviolet fluctuations are disproportionately responsible for solar irradiance variations, however, the true correction needed for visual wavelengths may be less than this factor suggests.

The photometric observations of the programme stars have a single-measurement nightly precision of 0.2–0.3%, depending mainly on the steadiness of the reference stars, and a year-to-year precision of $\sim 0.1\%$. The Sun's recent variability would be undetectable at this level of precision, so we were surprised to find that more than half of the 33 stars were clearly varying from year to year. The variations were typically 'smooth', either quasi-cyclic or monotonic, rather than stochastic.

Solar chromospheric emission, which may be detected conveniently by spectroscopic measurements in the core of the K line of ionized calcium, arises from the presence and evolution of magnetic fields on the Sun's surface. Although chromospheric emission originates mainly in localized regions on the Sun, the measurements used here are averages over the whole solar disk. Observations of full-disk solar chromospheric K-line emission since 1974 at the National Solar Observatory show that the cyclic magnetic activity of the Sun observed 'as a star' corresponds closely to the rise and fall of the sunspot number (refs 8, 9; and S. L. Keil, personal communication). Similar data, obtained since 1966 by the 'HK Project' at Mount Wilson Observatory, record chromospheric activity in the H and K lines of ionized calcium for almost 100 Sun-like stars, including the 33 stars observed photometrically at Lowell Observatory. Like the Sun, many of these stars have regular activity cycles; others, with lower average HK emission levels, are quiescent (non-cycling)^{10–12}. A definitive calibration between the slightly dissimilar measures of solar and stellar chromospheric activity remains to be done. Detailed reviews of available data by us and by others¹³ persuade us, however, that the uncertainty in

FIG. 1 *a*, Stellar photometric variability (b and y magnitudes averaged) as a function of the mean chromospheric emission ratio. Variability is measured by the standard deviation of the annual mean magnitude values corrected for measurement error and intrinsic variability of the comparison stars. Peak-to-peak amplitudes would be about three times larger. Filled circles denote stars deemed variable at the 95% confidence level or greater; the straight line segment represents the least squares fit to these data. Crosses denote the remainder, that is, stars that were not deemed variable with at least 95% confidence. The Sun's variation is based on the standard deviation of the annual mean values of total irradiance measured by the Solar Maximum Mission¹. The Sun's mean chromospheric emission ratio is based on analysis of solar K-line observations from the National Solar Observatory (W. C. Livingston, personal communication). *b*, Chromospheric emission variability, expressed as millionths of the total flux, derived for the stars from the Mount Wilson HK emission measurements¹⁷ and for the Sun from observations at the National Solar Observatory (W. C. Livingston, personal communication). The Sun's cyclic activity variation is comparable with other stars of similar mean activity. *c*, Slopes of the regression of brightness on the HK index for stars that showed a significant correlation between the two variables over the interval of parallel measurement, 1984–91. The solar observations were obtained during 1980–88. Because the Sun's recent irradiance variation was so small, it lies closer to the dividing line between direct and inverse correlation than any other star.



the conversion between solar and stellar scales is too small to affect our solar–stellar comparisons significantly.

The Lowell and Mount Wilson measurements, made in parallel since 1984, now permit a direct comparison of stellar brightness and magnetic activity changes with those of the Sun^{14–16}. We have selected the r.m.s. dispersion of annual mean values as our variability measure for two reasons: (1) the annual mean emphasizes long-term changes by suppressing short-time-scale effects such as rotational modulation, and (2) the r.m.s. dispersion should be less sensitive to the influence of outliers than other possible measures of variation such as peak-to-peak range.

Figure 1 summarizes our results. The abscissa of this figure indicates the fraction of a star's total luminous output appearing (on average) as chromospheric emission¹⁷. Mean chromospheric activity declines with stellar age¹⁸, so the abscissa is also an age sequence running from young stars (right) to old stars (left). The upper panel shows that the Sun's observed brightness variability between 1980 and 1988, corrected to stellar units and expressed as the r.m.s. dispersion of annual mean values, was only about one-quarter that typical of stars of comparable mean magnetic (or chromospheric) activity. Nearly all of the stars more active (that is, younger) than the Sun were demonstrably variable in the visual continuum at levels ranging from 0.1% (the smallest amount we can detect reliably) up to ~2.7% r.m.s. The stars whose mean activity (and hence, age) were closest to the Sun's varied at levels ranging between 0.1% and 0.4% r.m.s., with a (fitted) mean of ~0.16% r.m.s., whereas the Sun itself varied by only 0.04% r.m.s. In comparison, the middle panel of Fig. 1 shows that the Sun's chromospheric variability during the same time was actually slightly greater than that typical of stars of comparable mean activity.

The bottom panel of Fig. 1 shows the relationship between the variations in brightness and the variations in chromospheric activity. Combining the photometry from Lowell with the HK data from Mount Wilson, we computed the regression of the eight annual mean brightness values on the annual mean chromospheric HK emission levels for each star. Twenty stars showed a statistically significant correlation between the two quantities. For these stars, we have plotted the ratio of brightness variability to HK variability in the lower panel. There is a systematic trend running from the young, active stars on the right side of the diagram, where photometric brightness and chromospheric activity variations were anticorrelated, to the older, less active stars (like the Sun) on the left, where they

were directly correlated. Because the Sun's observed irradiance variation is so small, its location on the figure is quite close to the dividing line between direct and inverse correlation.

The data shown in Fig. 1 represent single cycles or less of solar and stellar activity. We know from the historic sunspot record that individual solar activity cycles differ considerably, so we are left with the suspicion that the vertical dispersion in Fig. 1 is only a snapshot of stellar variability. The dispersion does, however, suggest bounds for the locus of variation that any one star, including the present Sun, might map out over timescales of decades to centuries. The horizontal run of mean values provides a general idea of the very long-scale temporal evolution of photometric and chromospheric variation for a Sun-like star. Our observations argue that the Sun—at least recently—varied less from activity maximum to activity minimum than similar stars despite its relatively normal degree of chromospheric variation. This raises the question of whether the Sun's cyclic irradiance variation is always so small or whether larger changes occur in other cycles or over longer time intervals. □

Received 22 October; accepted 29 October 1992.

- Willson, R. C. & Hudson, H. S. *Nature* **351**, 42–44 (1991).
- Hoyt, D. V., Kyle, H. L., Hickey, J. R. & Maschhoff, R. H. *J. geophys. Res.* **97**, 51–63 (1992).
- Foukal, P. & Lean, J. *Science* **247**, 504–505 (1990).
- Lean, J., Skumanich, A. & White, O. *Geophys. Res. Lett.* **19**, 1591–1594 (1992).
- Zhang, Q., Soon, W. H., Baliunas, S. L., Lockwood, G. W. & Radick, R. R. *Nature* (submitted).
- Lockwood, G. W. & Skiff, B. A. *Luminosity Variations of Stars Similar to the Sun*, Air Force Geophysics Lab. Rep. AFGL-TR-88-0221 (Hanscomb AFB, 1988).
- Lockwood, G. W., Skiff, B. A. & Thompson, D. T. *Proc. Int. Astr. Un. Colloq. No. 136* (eds Young, A. & Elliott, I.) (in the press).
- White, O. R. & Livingston, W. C. *Astrophys. J.* **249**, 798–816 (1981).
- White, O. R. in *Solar Radiative Output Variation* (ed. Foukal, P.) 87–112 (National Center for Atmospheric Research, Boulder, 1987).
- Wilson, O. C. *Astrophys. J.* **226**, 379–386 (1978).
- Baliunas, S. L. & Vaughan, A. H. *Astr. Astrophys.* **23**, 379–412 (1985).
- Baliunas, S. L. & Jastrow, R. *Nature* **348**, 520–523 (1990).
- White, O. R., Skumanich, A., Lean, J., Livingston, W. C. & Keil, S. *Publ. astr. Soc. Pacif.* (in the press).
- Radick, R. R., Lockwood, G. W. & Baliunas, S. L. *Science* **247**, 39–44 (1990).
- Radick, R. R. in *The Sun in Time* (eds Sonnett, C. P., Giampapa, M. S. & Matthews, M. S.) 787–808 (Univ. of Arizona Press, Tucson, 1991).
- Radick, R. R. *Twelfth NSO/Sacramento Peak Summer Workshop—The Solar Cycle* (ed. Harvey, K.) *Astr. Soc. Pacif. Conf. Series* **27**, 450–464 (1992).
- Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K. & Vaughan, A. H. *Astrophys. J.* **279**, 763–769 (1984).
- Skumanich, A. *Astrophys. J.* **171**, 565–567 (1972).

ACKNOWLEDGEMENTS. We thank the Air Force Geophysics Laboratory and the Lowell Research Fund for support of the work at Lowell Observatory, and the Electric Power Research Institute, the Lounsbery Foundation, the American Petroleum Institute, the Mobil Foundation, and the Langley-Abbott and Scholarly Studies Fund of the Smithsonian Institution for support of the work at Mount Wilson Observatory.

A direct inversion method for surface structure determination from LEED intensities

Peijun Hu & D. A. King

Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

LOW-ENERGY electron diffraction (LEED) has become the most successful technique in surface crystallography¹, but because of the complexity of the surface-electron scattering interactions, analyses of LEED data are still conducted on a trial-and-error basis: a direct-inversion method for treating LEED intensity data remains an attractive goal². Building on recent theoretical and experimental developments in electron holography from surface structures^{3–16}, we show here that three-dimensional images with atomic resolution can be obtained by a direct transform of conventional LEED intensity spectra.

In LEED structural analyses, a collimated electron beam incident on a well-ordered single crystal surface under ultra-

high-vacuum conditions is backscattered into a series of Bragg beams. The intensity I of each diffracted beam is measured as a function of the incident electron beam energy, V , yielding a set of I - V spectra which form the basic input to the structural analysis. Direct Fourier transforms¹⁷ have proved^{18,19} ineffective because of strong multiple scattering. Analyses are therefore done by trial and error, involving guesses at the structure, calculation of intensity spectra including multiple scattering, and comparison with the experimental data.

Here, we first use a cluster approach²⁰ to show theoretically that LEED and electron holography have important similarities. The scattering geometry is schematically shown in the inset to Fig. 1a in which the vectors $\mathbf{a}$ and $\mathbf{b}$ represent the atom positions in a surface, and $\mathbf{R}$ is the vector from the origin to a detector. An incident electron wave singly scattered to the detector by atom $\mathbf{a}$, and another electron wave doubly scattered by atoms $\mathbf{a}$ and $\mathbf{b}$ before reaching the detector, can be written (refs 21, 22; P.H. and D.A.K., unpublished results)

$$\psi_{\mathbf{a}} = e^{i\mathbf{k}_{\text{in}} \cdot \mathbf{a} - i\mathbf{k}_{\mathbf{a}} \cdot \mathbf{R}} A_{\mathbf{a}}^{(1)} e^{i\mathbf{k}_{\mathbf{R}} \cdot \mathbf{R}} \quad (1)$$

$$\psi_{\mathbf{ab}} = e^{i\mathbf{k}_{\text{in}} \cdot \mathbf{a} - i\mathbf{k}_{\mathbf{a}} \cdot \mathbf{R}} A_{\mathbf{ab}}^{(2)} e^{i\mathbf{k}_{\mathbf{R}} \cdot \mathbf{R}} e^{i\mathbf{k}[(\mathbf{b}-\mathbf{a}) \cdot (\mathbf{b}-\mathbf{a}) \cdot \mathbf{R}]} \quad (2)$$

respectively, where the exponents represent the phases of the

scattered waves relative to the origin, $A_a^{(1)}$ and $A_{ab}^{(2)}$ are the scattering factors for singly and doubly scattered waves respectively, containing curved-wave effects. Higher-order scattered waves can be obtained in a similar form. According to the 'tree' scheme of ref. 20, all scattered waves in a multiple scattering 'tree' starting at atom **a** shown in Fig. 1a can be summed as

$$\psi_{a,tree} = e^{i\mathbf{k}_{in} \cdot \mathbf{a} - i\mathbf{k}_a \cdot \mathbf{R}} \times \left[A_a^{(1)} e^{i\mathbf{k}_R} + \sum_b (A_{ab}^{(2)} e^{i\mathbf{k}_R} e^{i\mathbf{k}[(\mathbf{b}-\mathbf{a})-(\mathbf{b}-\mathbf{a}) \cdot \hat{\mathbf{R}}]} + \dots) \right] \quad (3)$$

b being an atom positioned near atom **a** and multiple (>2) scattered waves being included in the term ' $\dots$ '. If a perfectly ordered surface structure is assumed, the position of the s th atom in the j th unit cell can be written¹

$$\mathbf{a} = \mathbf{a}_{js} = \mathbf{W}_j + \mathbf{w}_s = n\mathbf{u} + m\mathbf{v} + \mathbf{w}_s$$

where $\mathbf{W}_j$ is the vector of j th unit cell displaced from the origin by $\mathbf{W}_j$, $\mathbf{u}$ and $\mathbf{v}$ are the vectors of a unit cell, and $\mathbf{w}_s$ is the position of an atom relative to the unit cell origin. We denote all the terms in the square bracket as A_s , which are the same for every unit cell, and sum all the scattered waves illuminated by the incident plane wave within the probe coherence area, containing N^2 unit cells, to give

$$\begin{aligned} \psi_{total} &= \sum_j \sum_s \psi_{a_{js},tree} \\ &= \sum_n \sum_m e^{i(\mathbf{k}_{in} - \mathbf{k}_R) \cdot (n\mathbf{u} + m\mathbf{v})} \\ &\quad \times \sum_s e^{i(\mathbf{k}_{in} - \mathbf{k}_R) \cdot \mathbf{w}_s} A_s \end{aligned} \quad (4)$$

The first two summations are the structure factor^{1,22}, which amount to constant coefficients in the LEED I - V spectra,

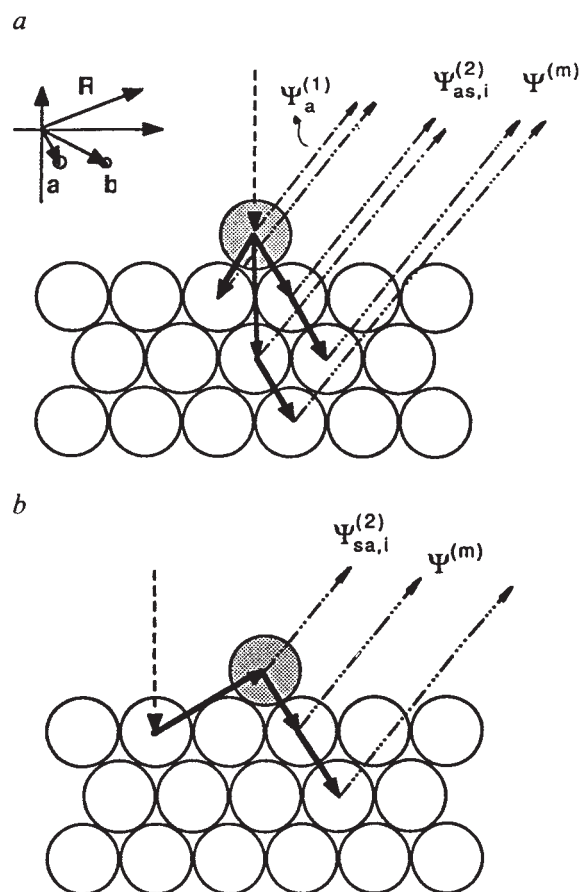
leading to

$$\psi_{total} \propto \sum_s e^{i(\mathbf{k}_{in} - \mathbf{k}_R) \cdot \mathbf{w}_s} A_s \quad (5)$$

Thus the main effect of long-range order, as distinct from short-range order, is to produce intensity at Bragg beam positions; it does not influence the I - V dependence at a given Bragg beam position. This result indicates that the LEED I - V spectra are dominated by scattering events occurring within a small region of crystal surface (ref. 22; P.H. and D.A.K., unpublished results), as in photoelectron or Auger electron diffraction. We now consider a particular case, a simple ordered overlayer surface. Diffracted beams from such a surface are conventionally divided into two categories: integral-order beams produced by the substrate alone, and fractional-order beams arising from the combined substrate and adsorbed layer symmetry. The fractional-order beams result from interference between scattered waves involving an adatom at least once. Equation (5) is then further simplified. Suppose that there is one adatom in a unit cell (Fig. 1). The scattered waves starting at the unit cell that end up in fractional-order beam positions can be ordered into two groups. The first group contains a wave singly scattered by the adatom, denoted as $\psi_a^{(1)}$, and all the waves doubly scattered first by the adatom and then by a substrate atom, denoted as $\sum \psi_{sa,i}^{(2)}$, where i is a substrate atom near the adatom, and the summation is over all the doubly scattered waves with sufficient large amplitudes. The second group consists of all the waves doubly scattered first by a substrate atom and then by the adatom, denoted as $\sum \psi_{sa,i}^{(2)}$, and all the multiply scattered waves involving the adatom at least once, denoted as $\sum \psi^{(m)}$, as shown in Fig. 1b. Waves singly scattered by substrate atoms and all the multiply scattered waves not involving the adatom are not included here, as they go to integral-order beam positions.

Using the terminology derived from holography, we can reinterpret these waves as follows. The wave singly scattered by the

FIG. 1 Schematic illustration of electron scattering in LEED. **a** Scattering 'tree' starting from an adsorbate atom. (---), incident wave; (---), wave singly scattered by the adatom, $\psi_a^{(1)}$; (---), waves scattered first by the adatom and then by substrate atoms, $\psi_{sa,i}^{(2)}$; (---), higher-order scattered waves, $\psi^{(m)}$. **b**, Multiple scattering starting from a substrate atom and then scattered by the adatom at least once. (---), waves scattered first by a substrate atom and then by the adatom, $\psi_{sa,i}^{(2)}$; (---) higher-order scattered waves, $\psi^{(m)}$. $\psi_a^{(1)}$ and $\psi_{sa,i}^{(2)}$ could be regarded as reference and object waves respectively. $\psi_{sa,i}^{(2)}$ and $\psi^{(m)}$ can be filtered out if a proper energy range of spectra is used in the transforms.



adatom is a reference wave, with its origin at the adatom, and the doubly scattered waves in the first group are the object waves. All the scattered waves in the second group can be filtered out, as shown later. The fractional-order beam intensities can be written, by analogy with holography, as

$$I = |\psi_a^{(1)}|^2 + |\sum \psi_{as,i}^{(2)} + \sum \psi_{sa,i}^{(2)} + \sum \psi^{(m)}|^2 + \psi_a^{(1)*} (\sum \psi_{as,i}^{(2)} + \sum \psi_{sa,i}^{(2)} + \sum \psi^{(m)}) + \psi_a^{(1)} (\sum \psi_{as,i}^{(2)*} + \sum \psi_{sa,i}^{(2)*} + \sum \psi^{(m)*}) \quad (6)$$

It can be shown that the first and second terms are actually the squares of amplitudes of scattered waves, and their major components have lower frequencies than do the third and the last terms, and can therefore be suppressed (refs 8, 14; P.H. and D.A.K., unpublished results). The third term is the real image term from which the object can be reconstructed and the last term is the twin image term, in holographic terminology. It should therefore be possible to use a direct holographic transform to generate a real-space image of a surface structure from LEED $I-V$ spectra, provided that fractional-order Bragg beams from a simple ordered overlayer structure are used. In principle, the twin image term could also be used to reconstruct the surface structure.

Barton⁴ proposed a formalism to reconstruct surface structures from photoelectron diffraction patterns, and scattering-factor corrections were later suggested^{11,12} for object waves. Multiple energy transformations can markedly improve the reconstructed images^{14,15} the filter concept in holographic transforms has also been addressed¹⁰. Based on this, we propose the following transform to obtain the structural information from LEED $I-V$ spectra:

$$U(\mathbf{r}) = \sum_k \sum_{\mathbf{R}} I \frac{e^{ik(-\mathbf{r}+\mathbf{r}\cdot\hat{\mathbf{R}})}}{A_a^{(1)*}(A_{as,\mathbf{r}}^{(2)}/|A_{as,\mathbf{r}}^{(2)}|)} \quad (7)$$

where

$$A_a^{(1)*} = \left[\frac{1}{ik} \sum_l (2l+1) T_{l,a} P_l(\cos \theta_{\mathbf{k}_{in},\mathbf{R}}) \right]^* \\ A_{as,\mathbf{r}}^{(2)} = -\frac{1}{k^2} \sum_l (2l+1) T_{l,a} P_l(\cos \theta_{\mathbf{k}_{in},\mathbf{r}}) d_l(kr) \\ \times \sum_{l'} (2l'+1) T_{l',s} P_{l'}(\cos \theta_{\mathbf{r},\mathbf{R}}) \\ d_l(kr) = \sum_{p=0}^l \frac{(1+p)!}{p!(l-p)!} \left(\frac{i}{2kr} \right)^p$$

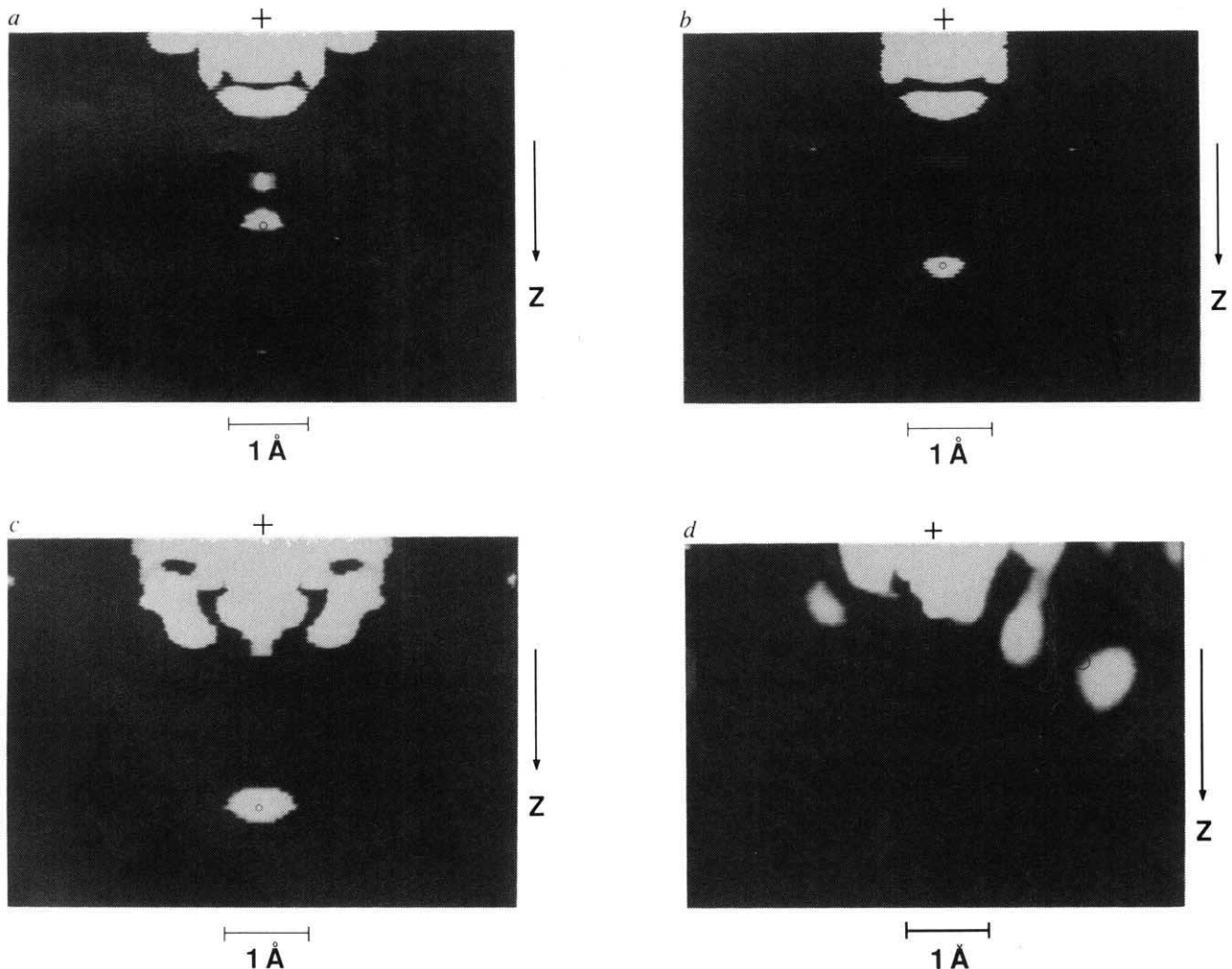


FIG. 2 Holographic images directly transformed from calculated fractional-order LEED $I-V$ spectra of Ni(100)- $p(2 \times 2)$ -O for O atoms in 4-fold hollow sites, (a) 0.8 Å, (b) 1.2 Å, (c) 1.6 Å above the first Ni layer respectively at normal electron beam incidence; d, 1.6 Å above the first Ni layer in off-normal incidence, viewed in a vertical cut along the $\langle 100 \rangle$ azimuth through the

O atom and Ni atoms. The plus signs are the O atom positions, and the centres of the circles are the Ni atom positions in the second layer for a, b and c, and in the first layer for d. A Ni atom directly below the O atom is well imaged because of anisotropies of electron scattering.

$T_l = (e^{i2\delta_l} - 1)/2$ and $P_l(\cos \theta)$ are Legendre polynomials. I denotes the fractional-order beam intensities from an ordered overlayer system. The summation over $\mathbf{R}$ in equation (7) refers to all available fractional-order Bragg beams at a given incidence energy, while the summation over k refers to the available energy or k range for which these intensities have been measured; the summations are used in place of the integrals used in treating intensity maps as holograms. The scattering-factor corrections for both the reference and the object waves $A_a^{(1)}$ and $A_{as,r}^{(2)}$ are also included. The origin is defined at the adatom position, the source of the reference wave. Three-dimensional images can be obtained by plotting intensities $|U(\mathbf{r})|^2$ with respect to a position $\mathbf{r}$ in real space.

How the transform works can be understood as follows. We consider the real image term first. For the term $\psi_a^{(1)*} \sum \psi_{as,i}^{(2)} = A_a^{(1)*} \sum A_{as}^{(2)} e^{ik(b_i - b_r) \cdot \mathbf{R}}$ obtained from equations (1) and (2), in the real image term of equation (6), if $\mathbf{r}$ is not equal to a substrate atom position $\mathbf{b}_i$ in a transform, the summations give rise to a small value because the scattering factors and the exponents are oscillating complex functions in the transform; when $\mathbf{r}$ is equal to $\mathbf{b}_i$, $e^{ik(-r+r) \cdot \mathbf{R}}$ and $(A_a^{(1)*} A_{as,r}^{(2)}) / |A_{as,r}^{(2)}|^{-1}$ tend to cancel out the scattering path length contributions and the scattering factors to unity, and summations should produce a large value. In the term $\psi_a^{(1)*} \sum \psi_{sa,i}^{(2)} = A_a^{(1)*} \sum A_{sa}^{(2)} e^{i(k_{in} b_i + k b) \cdot \mathbf{r}}$, and also in the real image term, the cancellation mentioned above where $\mathbf{r}$ is equal to an atom position cannot happen, and thus the transform should give rise to small values if a reasonably wide energy range (k range) is used: it can be filtered out. For the same reason, the multiple scattering term $\sum \psi^{(m)}$ is also filtered out. Multiple energy transformations can also suppress the first, the second and the twin terms in equation (6)¹⁴, which can be explained in the same way.

For a systematic test of this direct inversion method, we calculated three sets of LEED I - V spectra, including multiple scattering, for the system Ni{100}- $p(2 \times 2)$ -O at normal incidence. The oxygen adatoms were placed in 4-fold hollow sites on the square array of surface Ni atoms, with no relaxation of surface Ni atoms from bulk positions. The height of the O adatoms above the Ni surface plane was varied between 0.8 and 1.6 Å. Input parameters, such as phase shifts, were the same as those in a previous LEED calculation²³, which gives a good description of experimental data. In the transforms, the energy step $\Delta k = 0.07 \text{ Å}^{-1}$ is used, and the energy range is 30 eV to 273 eV.

Figure 2a-c shows images directly transformed from the LEED spectra of O atoms 0.8 Å, 1.2 Å and 1.6 Å, respectively, above the first Ni layer, viewed in a vertical cut through the O atom and Ni atoms. The plus signs are the O atom positions, the origin of the transformed structure in real space, and the centres of the small circles are the known Ni atom positions in the second layer. It is clear that the brightest spots (other than the origin) are always found at the correct Ni atom positions in the second Ni layer, whereas the Ni atom positions in the first layer are not clear. Electron scattering is anisotropic and the scattering amplitudes peak strongly in the forward direction. Therefore, it is not surprising that the atoms directly below the adatoms are well imaged. This can be clearly seen in Fig. 2d obtained by a direct transform of LEED I - V spectra in which a non-normal incident beam was directed along the axis between the O atom and one of Ni atoms in the first layer. In this case the first layer Ni atom is well imaged, but not the second layer Ni atom. As agreement between calculated and experimental I - V spectra is good, this method is clearly viable for the analysis of experimental data.

The primary limitation of the technique at present is that it applies only to simple overlayer adsorbate structures producing fractional-order LEED beams. But it has many advantages over other holographic transform methods⁴⁻¹⁵. Experimentally, good-quality LEED I - V spectra can be obtained at energies up to several hundred electronvolts without the difficulties associated

with photoelectron, Auger electron and diffuse LEED patterns. The incident angle can also be altered to increase the data base. In particular, this will increase the number of substrate atoms that can be imaged. Noise near the origin may be reduced by filtering out the low-frequency components of equations (6) (P.H. and D.A.K., unpublished results). We expect that increasing the energy range will also improve the image quality. □

Received 30 July; accepted 17 November 1992.

1. Pendry, J. B. *Low Energy Electron Diffraction* (Academic, London, 1974).
2. Pendry, J. B., Heinz, K. & Oed, W. *Phys. Rev. Lett.* **61**, 2953-2956 (1988).
3. Szöke, A. in *Short Wavelength Coherent Radiation: Generation and Application* (eds Attwood, D. T. & Boker, J.) *AIP Conf. No. 147* (American Institute of Physics, New York, 1986).
4. Barton, J. J. *Phys. Rev. Lett.* **61**, 1356-1359 (1988).
5. Saldin, D. K. & de Andres, P. L. *Phys. Rev. Lett.* **64**, 1270-1273 (1990).
6. Harp, G. R., Saldin, D. K. & Tonner, B. P. *Phys. Rev. Lett.* **65**, 1012-1015 (1990).
7. Tong, S. Y., Wei, C. M., Zhao, T. C., Huang, H. & Li, H. *Phys. Rev. Lett.* **66**, 60-63 (1991).
8. Hu, P., Barnes, C. J. & King, D. A. *Chem. Phys. Lett.* **183**, 521-528 (1991).
9. Hu, P. & King, D. A. *Nature* **353**, 831-833 (1991).
10. Hu, P. & King, D. A. *Phys. Rev. B* (in the press).
11. Tonner, B. P., Han, Z.-L., Harp, G. R. & Saldin, D. K. *Phys. Rev.* **B43**, 14423-14433 (1991).
12. Saldin, D. K., Harp, G. R., Chen, B. L. & Tonner, B. P. *Phys. Rev.* **B44**, 2480-2494 (1991).
13. Wei, C. M., Zhao, T. C. & Tong, S. Y. *Phys. Rev. Lett.* **65**, 2278-2280 (1990).
14. Barton, J. J. *Phys. Rev. Lett.* **67**, 3106-3109 (1991).
15. Tong, S. Y., Li, H. & Huang, H. *Phys. Rev. Lett.* **67**, 3102-3105 (1991).
16. Mendez, M. A., Gluck, C. & Heinz, K. *J. Phys.: Condens. Matter* **4**, 999 (1992).
17. Landman, U. & Adams, D. L. *Surf. Sci.* **51**, 149-173 (1975).
18. Marsh, F. S. & King, D. A. *Surf. Sci.* **79**, 445-460 (1979).
19. Woodruff, D. P. *Faraday Disc.* **60**, 218-229 (1975).
20. Barton, J. J., Xu, M.-L. & Van Hove, M. A. *Phys. Rev.* **B37**, 10475-10486 (1988).
21. Barton, J. J. & Shirley, D. A. *Phys. Rev.* **B32**, 1906-1920 (1985).
22. Hu, P., Barnes, C. J. & King, D. A. *Phys. Rev.* **B45**, 13595-13598 (1992).
23. Van Hove, M. A. & Tong, S. Y. *Surface Crystallography by LEED* (Springer, Berlin 1979).

ACKNOWLEDGEMENTS. We thank A. Wander for discussions on the dynamical LEED programme. The Oppenheimer Trust (UK) is acknowledged for a studentship to P.H.

Mediterranean vegetation, lake levels and palaeoclimate at the Last Glacial Maximum

I. Colin Prentice*, Joël Guiot† & Sandy P. Harrison‡

* Department of Plant Ecology, Lund University, Östra Vallgatan 14, S-22361 Lund, Sweden

† Laboratoire de Botanique Historique et Palynologie, CNRS UA1152, F-13397 Marseille Cedex 13, France

‡ Department of Earth Sciences, Uppsala University, Box 554, S-75122 Uppsala, Sweden

THE apparent conflict between pollen evidence for widespread *Artemisia* steppe¹⁻⁶ (implying semi-arid conditions) and geomorphological evidence for high lake levels^{4,7-11} has produced controversy about the ice-age palaeoclimate of the Mediterranean region. Here we use a water-balance model¹² (to predict catchment runoff) and a biome model¹³ (to predict vegetation type) to reconstruct the palaeoenvironment around Lake Ioannina—a type locality for the northern Mediterranean region. We show that both sets of evidence are compatible with a summer-dry, winter-wet regime with seasonal temperature anomalies similar to those predicted by atmospheric model simulations of the Last Glacial Maximum¹⁴⁻¹⁸. The drying effect of the cold North Atlantic Ocean may have been counteracted in winter by increased storm frequency under a southward-shifted jet stream, as shown by several atmospheric models¹⁶⁻¹⁸.

The occurrence of higher-than-present lake levels in *Artemisia*-dominated steppe during the millenia around the Last Glacial Maximum (~18 kyr BP) is a feature of palaeoenvironmental records in the northern Mediterranean region from Spain to Iran¹⁻¹¹. At Lake Ioannina (39° 40' N, 20° 53' E, 469 m above sea level (m a.s.l.)), terraces from a former high stand ≥3.2 m above the present lake level have been radiocarbon-dated at 20-21 kyr on hearths associated with human habitation of the shore^{7,8,11}. The pollen diagram shows dominance by the shrub

TABLE 1 Modern climate at Lake Ioannina, Greece

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Temperature (°C)	5.0	6.0	9.5	13.0	17.0	22.0	22.5	25.0	20.0	15.6	10.0	6.9
Precipitation (mm)	150	130	100	85	85	60	29	33	61	130	160	170
Sunshine (%)	43	44	48	59	68	74	86	82	73	61	45	34

genus *Artemisia* (30–40% of total pollen) together with grasses (15–25%) and Chenopodiaceae (~10%), with small and fluctuating amounts of tree pollen (mostly *Pinus* and *Quercus*; always <25%). This period lasted from ~30 kyr to ~15 kyr, on the basis of radiocarbon dates from the lake sediments^{2,3,11}.

The apparent conflict between dry conditions suggested by vegetation and wet conditions suggested by lakes may arise from the sensitivity of lake levels and vegetation to different components of the water balance. Overflowing lakes tend to an equilibrium level such that outflow balances the lake-surface water balance plus runoff from the catchment, and this level is sensitive to changes in runoff^{11,19}. Vegetation tends to an equilibrium foliage density such that evapotranspiration during wet periods conserves enough water for survival through dry periods²⁰; water that is not evaporated or transpired (for example, precipitation falling on already-saturated soils) accrues to runoff. Shifts in the seasonality of precipitation and potential evapotranspiration (PE) can therefore theoretically increase runoff at the expense of soil moisture, causing an increase in lake levels and a reduction in vegetation cover.

We apply these standard hydrological concepts to the environment of Lake Ioannina in an attempt to find a plausible palaeoclimatic scheme that could increase runoff substantially, while replacing the present vegetation by steppe. A soil water-balance model¹² is used to partition precipitation (P) into actual evapotranspiration (E) and runoff ($P - E$). From calculated insolation and monthly data on temperature, precipitation and sunshine, the model computes monthly values of PE as a function of net radiation and temperature and estimates E as a function of PE, maximum transpiration rate (here taken to be 1.0 mm h^{-1}) and soil wetness (water content as a fraction of total available water capacity, here taken to be 150 mm).

The ratio E/PE was used as an index of relative moisture availability to plants¹³. At a continental scale, water-limited vegetation boundaries are predicted closely by a few critical values of this index¹³. The biome model¹³ combines these with thermal limits to predict vegetation type. Broad-leaved evergreen trees and shrubs are assumed to require winter (mean coldest-month) temperatures of more than 5°C . Forests with broad-

leaved evergreen trees can occur where $E/PE > 0.65$. Sclerophyll woodlands and shrublands replace forest in drier climates with $E/PE < 0.65$. In climates with colder winters (-2°C to 5°C), temperate deciduous forests occur where $E/PE > 0.65$, and steppe (grasslands or dwarf shrublands dominated by *Artemisia* and/or Chenopodiaceae) where $E/PE < 0.65$. These limits accurately predict the modern European distributions of these biomes¹³.

The present natural vegetation near Lake Ioannina is Mediterranean forest with broad-leaved evergreen *Quercus coccifera*, *Q. ilex* and *Pistacia* spp. and deciduous *Q. pubescens* and *Carpinus orientalis*^{2,3}. This composition is reflected in the pollen assemblage from the surface lake sediments³. With monthly temperature and rainfall data from ref. 3 and sunshine data smoothly interpolated to the location and elevation of Ioannina from ref. 21 (Table 1), the water-balance model generated 496 mm of annual runoff (Fig. 1) and a moisture index (E/PE) of 0.69. Thus, winter temperatures are just warm enough to allow broad-leaved evergreens as well as temperate deciduous trees, and enough moisture is available for forests to survive.

In a series of sensitivity experiments (Table 2) we perturbed this modern climate on a seasonal basis, taking the six summer months to be those with more than 5 mm soil moisture deficit (May–October). Experiments 1–3 illustrate the effects of possible temperature anomalies for the glacial maximum. We considered decreases of up to 6 K in winter, in summer and all year round. These changes produced modest increases in runoff but no substantial reduction in the moisture index. In experiment 4, the winter cooling of experiment 1 accompanied by an excess of precipitation in the winter months produced a larger increase in runoff, again with no substantial change in the moisture index. But by combining this winter cooling with a redistribution of the same total annual precipitation through the year (with increases in winter balanced by decreases in summer: experiment 5), we were able to produce a large increase in runoff and a decrease in the moisture index to 0.57. A closely similar result was obtained in experiment 6 (Fig. 1), where a modest (2 K) reduction of summer temperatures was added to the situation of experiment 5.

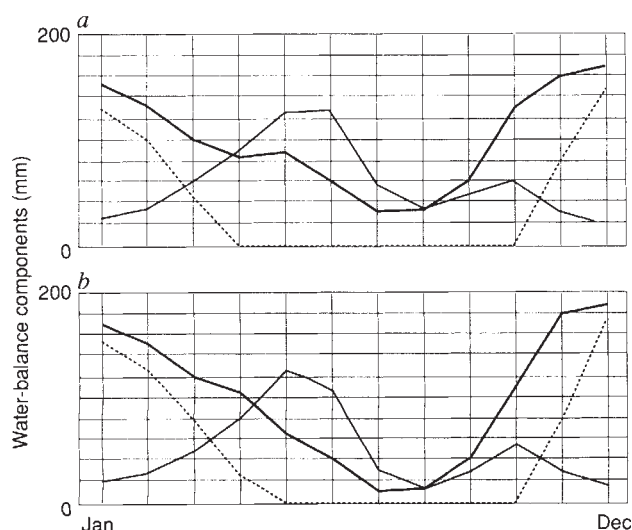


FIG. 1 Monthly precipitation (thick line), evapotranspiration (thin line) and runoff (dotted line) at Lake Ioannina. *a*, Modern simulation based on data in Table 1; *b*, experiment 6 with cold winters, cool summers and increased precipitation seasonality, described in Table 2. Monthly runoff is defined as $P - E - \delta W$ where P is monthly precipitation, E is monthly evapotranspiration, and δW is the change in soil moisture storage during the month.

The low value of the moisture index in experiments 5 and 6 arises because most of the excess precipitation falling in winter runs off from already-saturated soils, whereas a part of the shortfall in summer is compensated by reduced evapotranspiration from drier soils. The climates simulated in experiments 5 and 6 lie well within the climatic range for steppe because the winter temperature is too cold for evergreen shrubs, while moisture availability is inadequate for temperate deciduous trees.

It has alternatively been suggested²² that high lake levels might have been due to a combination of very cold winters and cool, cloudy summers. In experiment 7 we lowered winter temperatures by 10 K and summer temperatures by 2 K, and adjusted summer sunshine hours (by the method of ref. 12) to correspond to a 10° equatorward shift of the present circulation belts. Runoff increased less than in experiments 5 and 6, whereas the moisture index increased considerably. Winter was cold enough (-5°C) and E/PE sufficient (>0.75) for mixed (boreal coniferous/temperate deciduous) forest¹³. This result illustrates a general principle that simply lowering PE tends to reduce drought stress on vegetation as well as increasing runoff, in contrast with seasonality shifts where the same climate change can produce both an increase in drought stress and an increase in runoff.

We therefore suggest that the ice-age climate of the Mediterranean region was characterized by cold winters, intense winter precipitation and summer drought. This hypothesis differs from the conclusion that might be reached by inference from contemporary pollen analogues in Europe²³. The closest modern analogues in Europe for the full-glacial *Artemisia*-dominated pollen spectra are in the Ukraine, with low total annual precipitation and a spring or summer precipitation maximum^{23,24}. But *Artemisia* steppes with similar floristic dominance occur in regions with greater precipitation and an intense winter maximum, for example in the northern Great Basin of the United States, and provide an alternative vegetational and climatic analogue for the glacial Mediterranean steppes.

Experiments with atmospheric general circulation models (GCMs) forced with glacial-maximum ice sheets, sea-surface temperatures and (in some experiments) low atmospheric CO_2 levels¹⁴⁻¹⁸ indicate that temperatures in the Mediterranean region were 5–10 K lower than present in winter and 1–3 K lower than present in summer. Temperatures were reduced most in winter primarily because of strong westerly advection from the cold North Atlantic. Summer temperatures were only slightly reduced because of local heating and weaker advection. GCM simulations also show a year-round strengthening and equatorward shift of the jet stream across the North Atlantic¹⁴⁻¹⁸, and some show increased winter precipitation along the jet-stream axis extending far into the continent at or near the latitude of the Mediterranean¹⁶⁻¹⁸. This effect opposes a drying tendency due partly to the reduced moisture-carrying capacity of the winds over the cold ocean, and partly to the development of a fixed anticyclone over the North European ice sheet^{11,14}. The

TABLE 2 Simulated annual water-balance components at Ioannina

	P	PE	E	P-E	E/PE
Base case	1,193	1,009	697	496	0.69
Temperature changed					
1. Cold winter	1,193	974	665	528	0.68
2. Cold year-round	1,193	877	647	546	0.74
3. Cold summer	1,193	935	692	501	0.74
Temperature and precipitation changed					
4. Cold wet winter	1,313	974	665	648	0.68
5. Cold wet winter; dry summer	1,193	974	572	621	0.59
6. Cold wet winter; cool dry summer	1,193	946	570	623	0.60
Temperature and cloudiness changed					
7. Very cold winter; cool cloudy summer	1,193	793	612	581	0.77

Water-balance components are given in millimetres.

- 6 K cooler December–March; 3 K cooler April, November.
- 6 K cooler all months.
- 6 K cooler June–September; 3 K cooler May, October.
- As case 1 but 20 mm per month wetter November–April.
- As case 1 but 20 mm per month wetter November–April; 20 mm per month drier May–October.
- As case 5 but 2 K cooler May–October.
- 10 K cooler December–March; 5 K cooler April, November; 2 K cooler May–October; monthly sunshine 43, 44, 48, 53, 57, 58, 61, 58, 52, 48, 45, 34%.

extent to which this anticyclone penetrates the Mediterranean region varies among models, as does the precise latitude of the jet-stream axis, so the different parts of the Mediterranean region can appear wetter or drier in winter. In summer, the jet stream (although shifted well south of its present summer position) continued to track north of the Mediterranean, producing a dry summer climate in all models.

Our results point to increased seasonality of precipitation, indicated by some but not all GCM simulations, as a key to the full-glacial Mediterranean palaeoclimate. Our results agree with those GCM simulations¹⁶⁻¹⁸ in which the southward-shifted jet stream produced increased winter precipitation in the Mediterranean region. The excess precipitation contributed to runoff and was not available to compensate for the growing-season soil-moisture deficit, resulting in a winter-rain steppe environment unlike any found in Europe today. As the ice sheets receded, the jet-stream axis retreated northward, removing the source of additional winter precipitation while sea surface temperatures remained low¹⁴. This may explain why both lake levels^{10,11} and tree pollen percentages³ in the eastern Mediterranean region fell to a minimum during late glacial times, ~13 kyr ago. □

Received 24 April; accepted 27 October 1992.

- Bonatti, E. *Nature* **209**, 984–985 (1966).
- Bottema, S. *Proc. Prehist. Soc.* **33**, 26–29 (1967).
- Bottema, S. *Late Quaternary Vegetation History of Northwestern Greece* (Proefschrift, Groningen, 1974).
- Wright, H. E. Jr *Science* **194**, 385–388 (1976).
- van Zeist, W. & Bottema, S. in *Palaeoclimates, Palaeoenvironments and Human Communities in the Eastern Mediterranean region in Later Prehistory* (eds Bintliff, J. L. & van Zeist, W.) 277–321 (British Archaeological Reports, International Series no. 133, Oxford, 1982).
- Pons, A. & Reille, M. *Palaeogeogr. Palaeoclim. Palaeoecol.* **66**, 243–263 (1988).
- Higgs, E. S. & Vita-Finzi, C. *Proc. Prehist. Soc.* **32**, 1–29 (1966).
- Higgs, E. S., Vita-Finzi, C., Harris, D. R. & Fagg, A. E. *Proc. Prehist. Soc.* **33**, 1–26 (1967).
- Roberts, N. *Quat. Res.* **19**, 154–171 (1983).
- Wasylikowa, K. *Rev. Palaeobot. Palynol.* **2**, 313–318 (1967).
- Harrison, S. P. & Digerfeldt, G. *Quat. Sci. Rev.* (submitted).
- Harrison, S. P., Prentice, I. C. & Guiot, J. *Clim. Dynam.* (in the press).
- Prentice, I. C. et al. *J. Biogeogr.* **19**, 117–134 (1992).

- Harrison, S. P., Prentice, I. C. & Bartlein, P. J. *Quat. Sci. Rev.* **11**, 283–299 (1992).
- Kutzbach, J. E. & Guetter, P. J. *J. Atmos. Sci.* **43**, 1726–1759 (1986).
- Broccoli, A. J. & Manabe, S. *Clim. Dynam.* **1**, 87–99 (1987).
- Joussaume, S. thesis, Univ. Pierre et Marie Curie, Paris (1989).
- Lautenschlager, M. & Herterich, K. *J. geophys. Res.* **95**, 22547–22557 (1990).
- Street-Perrott, F. A. & Harrison, S. P. in *Paleoclimate Analysis and Modeling* (ed. Hecht, A. D.) 291–340 (Wiley, Chichester, 1985).
- Woodward, F. I. *Climate and Plant Distribution* (Cambridge Univ. Press, Cambridge, 1987).
- Leemans, R. & Cramer, W. *The IIASA Climate Database for Mean Monthly Values of Temperature, Precipitation and Cloudiness on a Terrestrial Grid* (IIASA, Laxenburg, 1991).
- COHMAP Members *Science* **241**, 1043–1052 (1988).
- Sabatier, R. thesis, Univ. des Sciences et Techniques du Languedoc (1983).
- Wijmstra, T. A., Young, R. & Witte, H. J. L. *Geol. Mijnb.* **69**, 243–251 (1990).

ACKNOWLEDGEMENTS. We thank S. Joussaume for discussions. This research was supported by the Swedish Natural Science Research Council (NFR) and the French National Research Council (CNRS) under a bilateral agreement.

Metasomatism of the sub-arc mantle inferred from trace elements in Philippine xenoliths

René C. Maury*, Marc J. Defant† & Jean-Louis Joron‡

* Laboratoire de Pétrologie et URA1278, Université de Bretagne Occidentale, 29287 Brest, France

† Department of Geology, University of South Florida, Tampa, Florida 33620, USA

‡ Laboratoire Pierre Süe, Centre d'Etudes Nucléaires de Saclay, B.P. 2, 91191 Gif sur Yvette, France

ISLAND arc basalts are thought to derive from the melting of the wedge of mantle overlying the subducting slab¹; hence, the composition of this mantle wedge, and the nature of any metasomatic fluids or magmas introduced from the slab, have been the subject of much speculation²⁻⁶. Most of the evidence bearing on these questions has been indirect, coming from the chemistry of island arc lavas, or from mantle xenoliths originating elsewhere than immediately below the arc^{7,8}. Here we report trace element data for mantle xenoliths from the Luzon arc (Philippines)⁹, which come from the mantle wedge directly below the arc-front volcanoes. The trace element patterns reflect metasomatism of a relatively depleted mantle (more depleted than the source of mid-ocean-ridge basalts), and identify the metasomatizing agent as a hydrous fluid, rather than a melt. We suggest that the depletion of high-field-strength elements in arc magmas originates from a combination of low initial concentrations in the mantle before metasomatism, and the inability of hydrous fluids to transport these elements.

The Luzon arc formed from the eastward subduction of the South China Sea oceanic lithosphere beneath the Philippines. Geochemical studies¹⁰⁻¹² have detected latitudinal variations along the strike of the arc which have been attributed to source enrichment through the contribution of subducted sediments of continental origin¹⁰⁻¹². Crustal blocks in the northern and southern termini of the arc are believed to be the source of these sediments. Batan Island is located in the northern section of the Luzon arc. Mantle xenoliths occur within high-K calc-alkaline basalts and andesites within pyroclastic-flow deposits (<0.1 Myr) from Mount Iraya, the youngest volcano on Batan.

Vidal *et al.*¹³ have shown that these xenoliths are enriched in incompatible elements. Their ⁸⁷Sr/⁸⁶Sr and ¹⁴³Nd/¹⁴⁴Nd ratios overlap with those of the host lavas despite the fact that the xenoliths have not equilibrated with them. The xenoliths are mainly spinel-bearing harzburgites, believed to have originated from the mantle wedge at a depth of about 30 km (ref. 14). Primary minerals (olivine, orthopyroxene and spinel) show evidence of high-pressure deformation. Variable amounts of phlogopite, sometimes associated with diopside, occur as late undeformed metasomatic phases. The xenoliths have been interpreted as previously metasomatized mantle fragments incorporated in the Luzon arc lavas during ascent. Thus, they do not represent the actual source of the host lavas, but their isotope characteristics indicate that they have undergone the same metasomatic event as the source of the host lavas¹³. Consequently, the Batan xenoliths indicate that metasomatism may affect even the upper sections of the mantle wedge by either (1) percolation of fluids/magmas well above the Benioff zone or (2) convection-induced dissemination of previously metasomatized parts of the mantle.

Major and trace element data from seven Batan peridotite xenoliths and a representative host lava are presented in Table

TABLE 1 Trace-element concentrations in Batan harzburgites and a representative host lava (B1)

	B1	B7a	B3f	B100	B101	B102	B110	B111
SiO ₂	55.0	45.75	46.00	44.85	42.60	43.80	41.40	44.83
TiO ₂	0.98	0.03	0.02	0.07	0.01	0.03	0.01	0.04
Al ₂ O ₃	18.05	1.91	1.82	1.25	1.20	0.88	0.60	1.20
Fe ₂ O ₃	7.70	8.58	8.16	8.59	9.20	8.95	8.78	8.46
MnO	0.17	0.15	0.13	0.13	0.14	0.14	0.16	0.11
MgO	3.66	41.54	40.08	43.35	44.10	43.00	48.00	43.11
CaO	8.30	1.61	1.32	0.09	0.19	0.85	0.20	1.33
Na ₂ O	3.40	0.15	0.12	0.06	0.02	0.07	0.03	0.03
K ₂ O	2.24	0.16	0.12	0.24	0.02	0.02	0.02	0.01
P ₂ O ₅	0.40	0.01	0.01	0.01	0.02	0.02	0.01	0.04
Total	99.90	99.89	97.53	98.64	97.50	97.76	99.21	99.16
Rb	112	10.6	2.7	13.3	0.86	1.49	1.45	1.34
Sr	653	34.9	8.71	—	—	—	—	5.91
Cs	8.35	10.61	0.143	0.597	0.052	0.081	0.083	0.092
Ba	677	39.6	15.8	27.3	4.9	5.5	8.1	5.0
La	38.3	1.26	0.443	0.280	0.508	0.273	0.187	0.463
Ce	85	4.0	3.2	—	2.58	—	—	0.60
Sm	6.06	0.22	0.068	0.066	0.043	0.046	—	0.067
Eu	1.69	0.089	0.0207	0.015	0.0183	0.052	0.051	0.021
Tb	0.52	0.0216	0.0106	0.0075	0.0082	0.0078	0.0010	0.0068
Yb	1.8	0.194	0.094	—	—	—	—	0.136
Hf	4.68	0.222	0.048	0.019	0.025	0.022	0.039	0.074
Ta	0.633	0.0244	0.0060	0.0014	0.0067	0.0057	0.0052	0.0061
Th	20.5	0.835	0.326	0.184	0.107	0.033	0.120	0.207
U	3.94	0.194	0.044	0.118	0.033	—	0.050	0.036
⁸⁷ Sr/ ⁸⁶ Sr	0.70479	0.70479	0.70471	—	—	—	—	0.70440
¹⁴³ Nd/ ¹⁴⁴ Nd	0.512506	0.512529	0.512451	—	—	—	—	0.512571

Trace element data (p.p.m.) were obtained by instrumental neutron activation analysis (INAA; C.E.N. Saclay, France) except for Rb and Sr (by isotope dilution in xenoliths¹³ and atomic absorption spectrometry in the B1 lava¹⁴). The neutron activation analyses were done in triplicate and average values are presented here. Mean analytical errors (in p.p.b.) are ± 6 (Th), ± 13 (Hf), ± 1.2 (Ta) and ± 2 (Tb). The accuracy for the other elements analysed by INAA is $\sim 5\%$. B1 is a high-K calc-alkaline basaltic andesite which is the host to xenoliths B100, B101, B102, B110 and B111. Decreasing amounts of metasomatic minerals (phlogopite, diopside for B7a and phlogopite for B3f and B100) are found from B7a through B3f to B100. The other four xenoliths lack modal metasomatic phases. Isotopic data is from Vidal *et al.*¹³. B7a has been labelled a lherzolite in this paper¹³.

1. We have excluded from this study those xenoliths which display mineralogical indications of interaction with the host lavas (for example, the occurrence of pargasite or Ti-augite—there are pargasitic reaction rims around the xenoliths)¹³. The data set (Table 1) contains some of the lowest Ta concentrations ever reported, similar to those from residual spinel peridotite xenoliths from the continental mantle¹⁵. The results are displayed on a chondrite-normalized diagram in Fig. 1a. The concentrations are very low for the high-field-strength elements (HFSE, such as Hf, Ta and Ti) and the middle and heavy rare-earth elements (MREE and HREE, respectively): they are typical of residual mantle peridotites^{15,16}. In contrast, Cs, Rb, Ba and K as well as the less incompatible elements Th, U, La and Ce are enriched relative to the former elements. Such trends have been interpreted as resulting from metasomatism affecting previously depleted mantle^{13,17,18}. The general shape of the xenolith trends is roughly similar to that of the host lava, and the most distinctive characteristic of these trends is the strong negative Ta anomalies typical of orogenic magmas⁴¹. The scatter observed among the xenoliths may reflect both their variable enrichment, due to metasomatism, and their variable refractory nature before metasomatism. The most incompatible-element-enriched xenoliths (B7a, B3f and B100) contain late undeformed phlogopite and thus show some evidence of modal metasomatism. These minerals have not been detected in the other four peridotites. Enrichment of the latter xenoliths can be ascribed to 'cryptic' metasomatism¹³.

Typical mantle compositions from the literature have been plotted on Fig. 1b along with two xenoliths (B7a and B3f). The higher concentrations of Cs, Rb, Ba, Th, U, K and perhaps La and Ce in the xenoliths compared with the published mantle compositions (Fig. 1b) emphasizes the important role of metasomatism. The HFSE, MREE and HREE contents may be used to determine the composition of the pre-metasomatized mantle (PMM) if we assume that these elements were not carried by the metasomatic fluids/magmas. Figure 1b shows that the concentrations of these elements, including Ta, are lower than Wood's¹⁹ values for enriched mantle (containing an ocean-island basalt (OIB) component) as well as the primitive mantles²⁻²². In fact, most of these elements fall below the values of ref. 19 and even ref. 23 for depleted MORB mantle.

There seems to be a consensus among arc petrologists that the PMM is residual mantle similar to the source of normal mid-ocean-ridge basalts (N-type MORB mantle)^{24,25} or even more refractory because of successive partial melting episodes^{6,26}. PMM similar in composition to the source of OIB, however, has been considered^{4,27}. Figure 1b seems to support the contention that the island-arc PMM is analogous to or more refractory than depleted MORB mantle (at least the section of the sub-arc mantle that was sampled by the magmas), because of successive melting events. Although our study cannot rule out the possibility that there are pods of OIB mantle below the Luzon arc that have not been sampled, the similarities between the Luzon arc lavas and most island-arc rocks (at least in terms of trace elements) suggest that an OIB source is not required to explain the geochemistry of arc magmas.

Many authors have postulated that negative Nb and Ta anomalies in island-arc magmas result from the stabilization of titanate phases in the mantle source^{4,28,29}. Theoretical³⁰ and experimental work²⁶ shows that a residual titanate phase could not exist in the sub-arc mantle. It is difficult to see how an OIB-type mantle could exist below arcs without a stable titanate phase in the source to produce the HFSE anomalies⁴. It has also been proposed that the HFSE anomalies may be generated by mantle-magma interactions³¹ or may even exist as a general feature of the mantle³². In contrast, our data support the contention that low concentrations of Ta are an inherent property of the arc-source region. The negative Ta anomalies shown in Fig. 1b can be explained by adding large-ion lithophile elements (LILE), Th, U and LREE by means of metasomatism to a

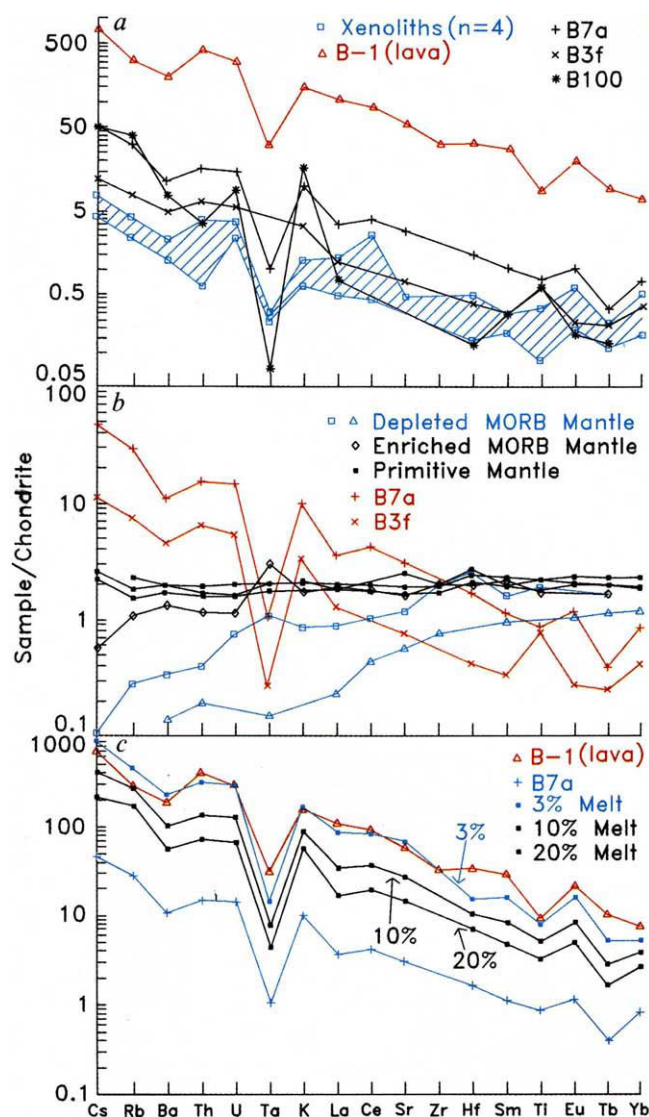
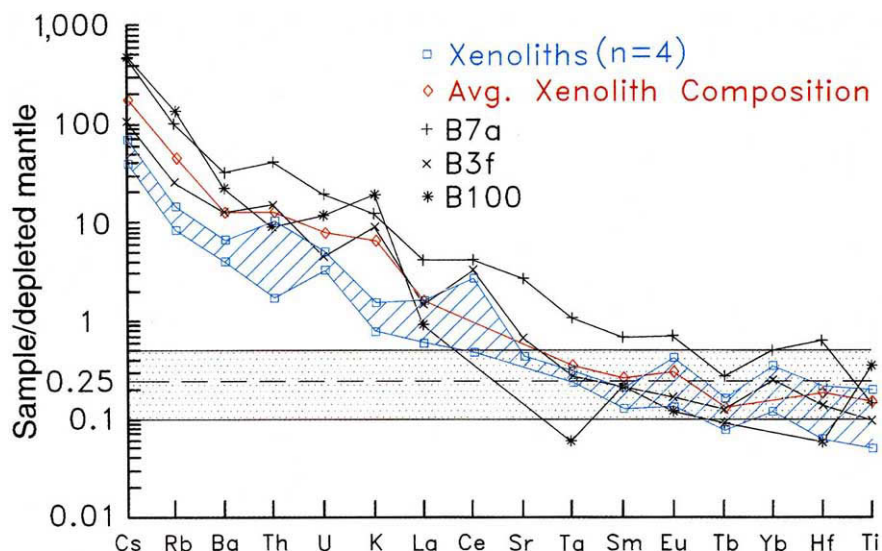


FIG. 1 a, Chondrite-normalized diagram for incompatible elements arranged according to their degree of incompatibility³⁷⁻³⁹. The stippled field delineates the range observed for the four most depleted xenoliths. b, Chondrite-normalized diagram similar to a showing various mantle compositions: depleted MORB mantle²³ (Δ), depleted ($\square$), and enriched ($\diamond$) MORB mantles¹⁹, primitive mantles ($\blacksquare$)^{20,21}, and mantle composition derived from alkali basalt ultramafic nodules ($\blacksquare$)²². Two xenoliths, B7a (+) and B3f ($\times$), are also included on the diagram. c, Chondrite-normalized diagram of a host lava B1 (Δ), phlogopite-bearing harzburgitic xenolith B7a (+), and curves generated from modal batch melting of this xenolith. The initial proportions of the minerals in the source rock¹⁴ are: olivine 0.606, orthopyroxene 0.245, clinopyroxene 0.021, phlogopite 0.11 and Cr spinel 0.018. Trace element partition coefficients are taken from refs 1 and 40.

depleted MORB-like PMM³³. The incompatible-element spectrum of the host lava can be adequately reproduced (Fig. 1c) through batch partial melting (20%) of the most incompatible-element-enriched peridotite (B7a). This modelling qualitatively supports our assumption that the xenoliths are geochemically similar to the source of the host lavas and that low Ta concentrations in the lavas are inherited from the mantle (not from the presence of stable titanate phases). It should be stressed, however, that the xenoliths are harzburgites and are therefore probably not the source of the basaltic parents, as some clinopyroxene is usually required in the source to produce basalt. The harzburgitic composition of the xenoliths can easily be explained, however, by varying degrees of depletion in the source region.

FIG. 2 Graph of xenolith compositions normalized to depleted MORB mantle¹⁹. Elements are given in order of increasing enrichment, from right to left. $\diamond$, Average xenolith composition. Stippled region equivalent to depleted MORB mantle before metasomatism (PMM).



To assess element mobilities during metasomatism, we have normalized the xenolith compositions to Wood's depleted MORB mantle¹⁹, and the elements are plotted in Fig. 2 in order of increasing enrichment. The mean composition of the seven xenoliths analysed has been included. Most of the concentrations of the elements from Ti to Ta on Fig. 2 fall within the stippled field which corresponds to 0.1 to 0.5 times Wood's depleted mantle (that is, these elements are typical of a depleted MORB mantle). Metasomatic enrichment may be evaluated for elements from Cs to the light rare-earth elements (LREE) using an approximated PMM composition that is 4 times more depleted than N-type MORB mantle (dashed line on Fig. 2). The observed enrichment varies widely from one xenolith to the other, depending on the presence or absence of metasomatic minerals. Typical enrichment factors obtained from the mean composition are: $\times 700$ (Cs), $\times 180$ (Rb), $\times 50$ (Ba and Th), $\times 30$ (U), $\times 25$ (K) and $\times 5$ (La).

Experiments²⁴ have shown that the mobility of incompatible elements in metasomatic fluids increases with ionic radii. The order of enrichment shown in Fig. 2 is roughly equivalent to increasing ionic radii except for K which appears less enriched than Th and U. Two out of the three phlogopite-bearing xenoliths, however, display positive K anomalies in Fig. 2 suggesting that the proposed order of the elements is sensitive to the type of metasomatism (cryptic against modal). The data displayed in Fig. 2 seem to be compatible with fluid-induced metasomatic enrichment of the Batan xenoliths. Ryerson and Watson²⁶ have demonstrated that negative Ta-Nb-Ti anomalies may alternatively be introduced into the mantle wedge by silicic magmas resulting from the melting of the subducted lithosphere under hydrous conditions. Melting of the subducted slab has been documented where young crust is subducted^{34,35}. The interaction of these slab-derived melts with mantle peridotites should produce enrichments in Sr, Na, Al and Eu in the mantle which are not detected in the Batan xenoliths. We therefore suggest that the metasomatizing agent is primarily fluid rather than silicic magma.

In relating the composition of the Batan xenoliths to the metasomatism involved in the generation of the arc magmas, one should remember that the xenoliths are known (from geobarometry¹⁴) to come from depths well above the zone of melt generation. In traversing the ~ 60 – 70 km between the subducting slab and the depth of origin of the xenoliths, the metasomatic fluids will almost certainly have reacted with the surrounding mantle, and may have changed in composition owing to chromatographic effects³⁶. But the agreement between the order of increasing enrichment of the trace elements in the xenoliths and Tatsumi's experimental results suggests that the fluid

composition did not change very much during transport.

Finally, we cannot rule out the possibility that the mantle sampled by the xenoliths may have been metasomatized by events unrelated to the Luzon arc. But the fact that our modelling shows that the Luzon lavas can be derived from the xenoliths (plus clinopyroxene), and that the xenoliths have the same unusual isotopic characteristics as the lavas (both groups fall below the mantle array on a radiogenic strontium versus neodymium diagram) suggests that the xenoliths have undergone metasomatism similar to that experienced by the source of the lavas. $\square$

Received 14 April; accepted 27 October 1992.

- Gill, J. B. *Organic Andesites and Plate Tectonics* (Springer, Berlin, 1981).
- Pearce, J. A. & Norry, M. J. *Contrib. Mineral. Petrol.* **69**, 33–47 (1979).
- Kay, R. W. *J. Geol.* **88**, 497–522 (1980).
- Morris, J. D. & Hart, S. R. *Geochim. cosmochim. Acta* **47**, 2015–2030 (1983).
- Perfit, M. R. & Kay, R. W. *Geochim. cosmochim. Acta* **50**, 477–481 (1986).
- Davidson, J. P. *Geochim. cosmochim. Acta* **51**, 2185–2198 (1987).
- Tanaka, T. & Aoki, K. *J. Geol.* **89**, 369–390 (1981).
- Takahashi, E. *J. Volc. geotherm. Res.* **29**, 355–395 (1986).
- Richard, M. et al. *Philippine Bull. Volc.* **3**, 1–27 (1986).
- Knittel, U. & Defant, M. J. *Earth planet. Sci. Lett.* **87**, 87–99 (1988).
- Defant, M. J. et al. *Tectonophysics* **183**, 187–205 (1990).
- Defant, M. J., Maury, R. C., Ripley, E. M., Feigenson, M. D. & Jacques, D. *J. Petrol.* **32**, 455–500 (1991).
- Vidal, P., Dupuy, C., Maury, R. C. & Richard, M. *Geology* **17**, 1115–1118 (1989).
- Richard, M. thesis, Bretagne Occidentale Univ. (1986).
- Jochum, K. P., McDonough, W. F., Palme, H. & Spettel, B. *Nature* **340**, 548–550 (1989).
- McDonough, W. F. & Frey, F. A. in *Geochemistry and Mineralogy of Rare Earth Elements* (eds Lipin, B. R. & McKay, G. A.) 99–145 (Miner. Soc. Am., 1989).
- Stosch, H. G. & Seck, H. A. *Geochim. cosmochim. Acta* **44**, 457–470 (1980).
- Menzies, M. A., Rogers, N., Tindle, A. & Hawkesworth, C. J. in *Mantle Metasomatism* (eds Menzies, M. A. & Hawkesworth, C. J.) 313–361 (Academic, London, 1987).
- Wood, D. A. *Geology* **7**, 499–503 (1979).
- Hofmann, A. W. *Earth planet. Sci. Lett.* **90**, 297–314 (1988).
- Sun, S. S. & McDonough, W. F. in *Magmatism in the Ocean Basins* (eds Saunders, A. D. & Norry, M. J.) 313–345 (Geol. Soc. Lond. Spec. Publ. **42**, 1989).
- Jagoutz, E. et al. *Proc. 10th lunar planet. Sci. Conf.*, 2031–2050 (1979).
- Hole, M. J., Saunders, A. D., Marriner, G. F. & Tarney, J. *J. geol. Soc. Lond.* **141**, 453–472 (1984).
- Tatsumi, Y., Hamilton, D. L. & Nesbitt, R. W. *J. Volcan. geotherm. Res.* **29**, 293–309 (1986).
- Nohda, S. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **52**, 265–276 (1981).
- Ryerson, F. J. & Watson, E. B. *Earth planet. Sci. Lett.* **86**, 225–239 (1987).
- Reagan, M. K. & Gill, J. B. *J. geophys. Res.* **94**, 4619–4633 (1989).
- Green, T. H. *J. Volcan. geotherm. Res.* **10**, 405–422 (1981).
- Foley, S. F. & Wheller, G. E. *Chem. Geol.* **85**, 1–18 (1990).
- Arculus, R. J. & Powell, R. *J. geophys. Res.* **91**, 5913–5926 (1986).
- Kelemen, P. B., Johnson, K. T. M., Kinzler, R. J. & Irving, A. J. *Nature* **345**, 521–524 (1990).
- Salter, V. J. M. & Shimizu, N. *Geochim. cosmochim. Acta* **52**, 2177–2182 (1988).
- McClulloch, M. T. & Gamble, J. A. *Earth planet. Sci. Lett.* **102**, 358–374 (1991).
- Drummond, M. S. & Defant, M. J. *J. geophys. Res.* **95**, 21503–21521 (1990).
- Defant, M. J. & Drummond, M. S. *Nature* **347**, 662–665 (1990).
- Navon, O. & Stolper, E. *J. Geol.* **95**, 285–307 (1987).
- Wakita, H., Rey, P. & Schmitt, R. A. *Proc. 2nd lunar Sci. Conf.*, 1319–1329 (1971).
- Sun, S. S., Nesbitt, R. W. & Sharaskin, A. Y. *Earth planet. Sci. Lett.* **44**, 119–138 (1979).
- Sun, S. S. & Nesbitt, R. W. *Earth planet. Sci. Lett.* **35**, 429–448 (1977).
- Henderson, P. *Inorganic Geochemistry* (Pergamon, Oxford, 1982).
- Joron, J. L. & Treuil, M. *Bull. Soc. Geol. France* **19**, 1197–1205 (1977).

ACKNOWLEDGEMENTS. We thank R. J. Arculus, C. Dupuy, C. Hawkesworth, A. Hochstaedter, P. Kepezhinskas, J. Ryan and H. Staudigel for comments on earlier versions of this manuscript, and M. Bordelo for drafting.

Crumbling of dacite dome lava and generation of pyroclastic flows at Unzen volcano

Hiroaki Sato*, Toshitsugu Fujii† & Setsuya Nakada‡

* Department of Natural Environmental Sciences, Faculty of Integrated Arts and Sciences, Hiroshima University, Hiroshima, 730 Japan

† Earthquake Research Institute, University of Tokyo, Bunkyo-ku, Tokyo, 113 Japan

‡ Department of Earth and Planetary Sciences, Faculty of Science, Kyushu University, Hakozaki, Fukuoka, 812 Japan

RECENT modelling of volcanic eruptions has shown that the efficiency of subsurface degassing of magmas determines whether magma erupts explosively or effuses quietly^{1,2}. Slow uprise of magma is often accompanied by effective degassing, leading to the extrusion of lava flows and domes. Although lava dome extrusion is one of the less explosive modes of eruption, it is often accompanied by explosive pyroclastic activities^{3–5}. The 1991 eruption of Unzen volcano provided an opportunity to observe at close range

several types of small-scale pyroclastic flow (glowing avalanches) originating from lava domes. Most of the pyroclastic flows are of Merapi type, caused by blocks falling from a collapsing dome; others are of Peléan type originating in an explosion from the side of a dome. The lavas apparently show variable degrees of self-explosivity. We suggest that variable degrees of degassing of the magma produced a wide range of excess pore pressures in the extruded lava domes, resulting in both Merapi-type and Peléan-type pyroclastic flows from the domes.

The extrusion of viscous lavas near the summit edge of the Fugendake dome of Unzen volcano, western Japan, began on 20 May 1991, and has been continuing at a rate of $1\text{--}3 \times 10^5 \text{ m}^3$ per day for more than a year^{6–8}. About 4,000 small-scale pyroclastic flows from the domes have been recorded seismically in the first year. One of the largest pyroclastic flows, on 3 June 1991, killed 43 people including three volcanologists⁹.

Close examination of videotapes verifies that instantaneous explosive fragmentation of a falling block causes a Merapi-type pyroclastic flow. Figure 1 illustrates the mode of collapse of a block falling from the second lava dome on 8 August 1991. The falling block is $\sim 70 \text{ m}$ high. The pyroclastic flow that accompanied this collapse was one of the smallest of the eruption, with a tremor lasting only 30 s, and was also the least explosive.

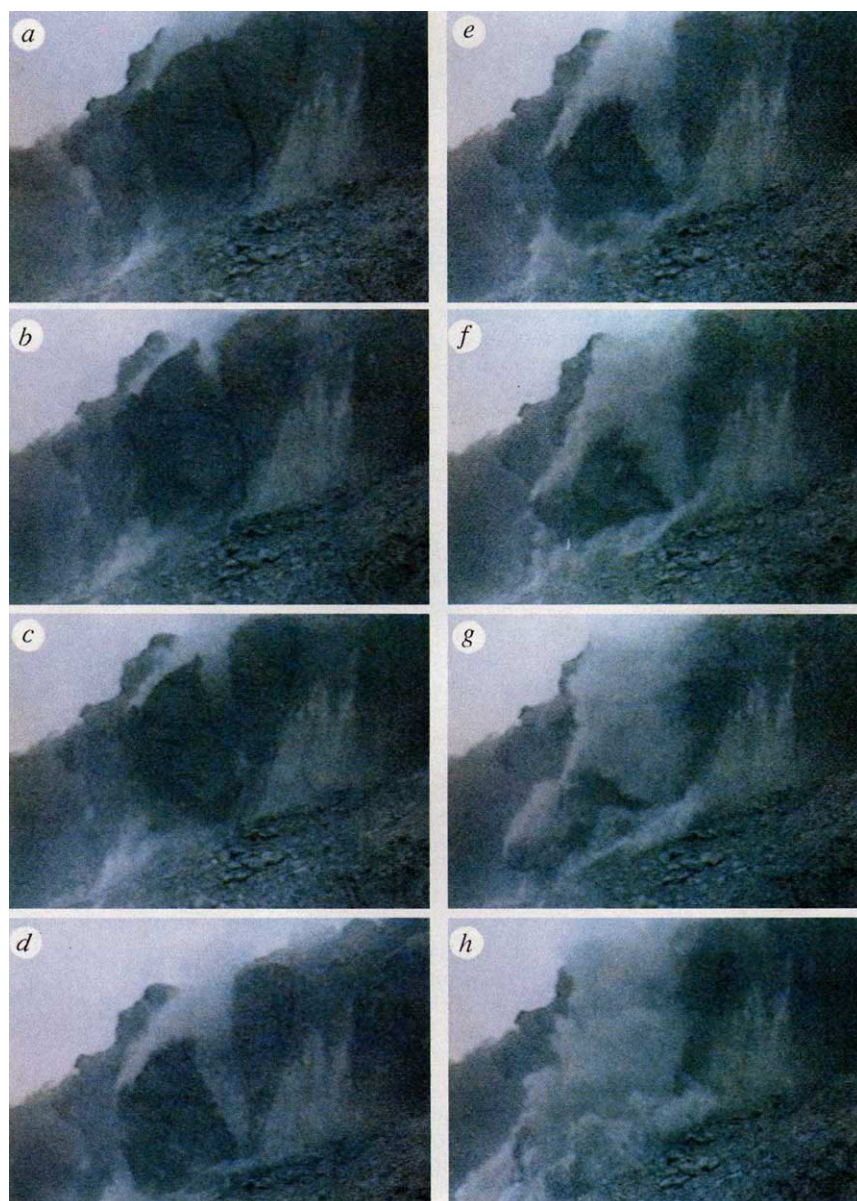


FIG. 1 Mode of crumbling of a lava block at the front of the second lava dome of Unzen volcano, at 16:17 on 8 August 1991. The block is $\sim 70 \text{ m}$ high. The photograph is copied from a videotape (NHK Japan) obtained $\sim 5 \text{ km}$ northeast of the lava dome. Frames are at 1-s intervals.

Ash-charged gas is issued from the surface of the block as soon as it begins to fall, then the block crumbles from below by explosive fragmentation. The trigger for fragmentation may be unloading¹⁰ or landing shock, or both. The process resembles the explosive demolition of a building. Similar but less definitive pictures were obtained during the 1980–86 eruption of Mount St Helens¹⁰, which involved dome building and collapse. The process implies that the falling block has low tensile strength, and some internal pore-gas pressure.

Self *et al.*¹¹ suggested that the likely range of tensile strength of andesite at subsolidus temperature is 100–200 bar. Dacite lava from the 1991 Unzen eruption, however, is highly viscous and contains 30–40 vol% vesicles, so it may have lower tensile strength. The lava samples generally consist of 30–50 vol% of crystals (phenocrysts and microlites) and glass. Lower tensile strength could be caused by high strain in viscous melt due to flow of the lava and presence of crystals. The matrix glass is rhyolitic in composition, and the viscosity (calculated after the method of Shaw¹²) is 10^{12} poise (water-free) to 10^9 poise (for 2 wt% of water) at 800 °C. Although we do not have experimental data on the tensile strength of the lava, field examination of hot blocks indicated it to be low: striking a block in a pyroclastic flow deposit with a hammer shattered it into fine particles. A similar observation was made at Mount St Helens¹⁰.

The explosive degassing of the lava blocks suggests that the dome lava contained high-pressure pore gas. The high concentration of volatiles in silicic domes leading to explosive collapse was proposed by Rose *et al.*³ and amplified by Fink *et al.*^{5,8,13}. The water content in the matrix glass may indicate the pore-gas pressure. Maeda *et al.*¹⁴ reported bulk-rock $H_2O(+)$ contents for the lava blocks in the range 0.18–0.46 wt%. Estimating the water content of the matrix glass is difficult because of uncertainty about the water content of matrix pargasite. The H_2O content of the bulk matrix is calculated to be 0.10–0.45 wt%, corresponding to an equilibrium saturation pressure of ~1–13 bar. These are probably only minimal values of the pore pressure of the dome lavas, because the analyses were made on lava blocks in pyroclastic flows, possibly the least explosive part of the dome lava, and because water could be degassed after the fragmentation.

High pore-gas pressure is suggested by the explosive growth of the ash clouds arising from pyroclastic flows. Air entrainment¹⁵ apparently did not take place near the front of these pyroclastic flows. Rather, the ash-charged clouds expanded upwards and forwards from the base of the advancing front of pyroclastic flows. This observation suggests that shearing at the bottom of the pyroclastic flow breaks large blocks into fine particles, where high-pressure pore gas is liberated and expands, charged with ash. The thickness of the cloud varies from one pyroclastic flow to another. Pyroclastic flows generated from new lava blocks near the vent are much more explosive and have thick billowing clouds, whereas those from the lava front at some distance from the source vent (Fig. 1) are generally less explosive. Because of its high viscosity, the lava resists the expansive force of bubbles during depressurization, so it is likely that the lavas just extruded from the vent have excess pore pressure (over hydrostatic pressure). Sparks¹⁶ simulated bubble formation numerically, and pointed out that viscosities of the order of 10^8 poise can cause excess pore pressures of several tens of bars. The apparent extrusion rate (several tens of metres per day) and high viscosity ($5\text{--}13 \times 10^{10}$ poise)¹⁷ suggest that the lavas just extruded could have excess pore pressures of tens of bars if degassing of water can be neglected. The excess pore pressure may decrease towards the surface of the lava because of cooling, and because jointing and fracturing of the lava surface caused by flow release the inner gas. Pyroclastic flows are often preceded by non-explosive block falls. Possibly, a fall of cool surface blocks is followed by explosive collapse of inner hot lavas to generate pyroclastic flows.

The mode of eruption of pyroclastic flows from lava domes,

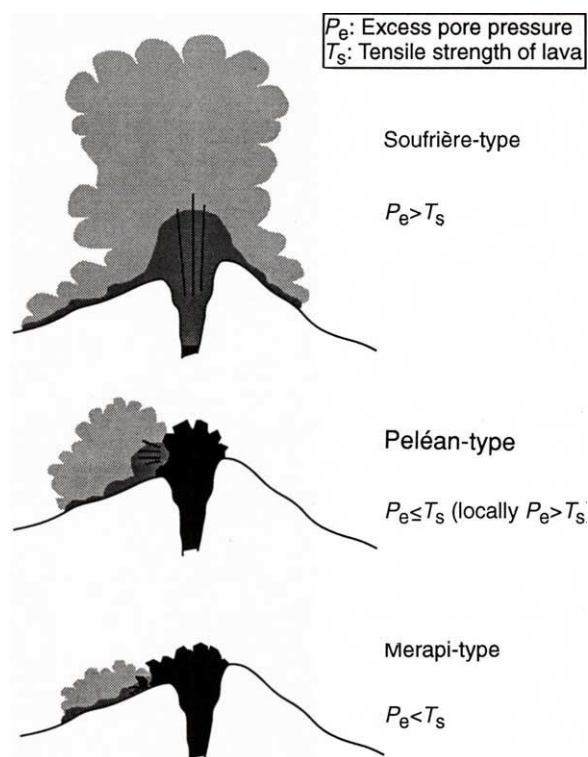


FIG. 2 Model of generation of pyroclastic flows in terms of excess pore pressure and tensile strength of lavas. Black area, massive lava; densely dotted area, densely charged pyroclastics and block and ash flows; lightly dotted area, ash-charged cloud.

in this case mainly Peléan and Merapi types, can be interpreted in terms of the relationship between tensile strength and excess pore pressure of lavas (Fig. 2). If the tensile strength of dome lava exceeds the excess pore pressure, the dome does not disintegrate unless a block fall triggers fragmentation. On the other hand, lateral explosion in the Peléan-type pyroclastic flow is generated where a part of the extruded dome lava has pore pressure greater than the tensile strength of the lava. Extruded dome lava is probably heterogeneous in water content and excess pore pressure (water distribution in the Mount St Helens dome was heterogeneous¹⁸). The initial water content of the 1991 Unzen lava is estimated from high-pressure phase equilibrium studies¹⁹ to be more than 4 wt%. The dome lava lost most of this water before reaching the surface. Details of the degassing mechanisms of viscous lavas are controversial^{2,13,20}, but it is conceivable that forced flow of the lava in the conduit makes cracks between pores in the lava, which may subsequently be annealed, but give some gas permeability. This is similar to percolation processes, which often show fractal size distribution. If the lava is similarly degassed by channelling of the pores, the water in the lava may be heterogeneously distributed. The pore size in bread-crust bombs ejected in the 11 June 1991 Vulcanian explosion of the Unzen lava dome ranges from less than $10\text{ }\mu\text{m}$ up to tens of centimetres, with no characteristic size. The bombs also show heterogeneous porosity distribution. Therefore, excess pore pressure is irregularly distributed in the extruded dome lavas, and Peléan-type pyroclastic flow may occur where excess pore pressure of the newly extruded lava locally exceeds the tensile strength of the lava.

If all of the dome lava had excess pore pressure greater than the tensile strength, the dome would explode instantaneously; this may correspond either to a Vulcanian-type explosion or a Soufrière- (or St Vincent-) type pyroclastic flow eruption^{21,22} (Fig. 2). During the 1991–92 eruption of Unzen volcano, Vulcanian-type explosions took place on 8 June and again on 11

June 1991. The 8 June event was preceded by seven Merapi-type and Peléan-type pyroclastic flows, which removed the capping dome of the lava conduit. Rapid removal of the cap released the pressure of the underlying water-rich magmas, which had higher excess pore pressure, causing an explosive eruption.

We thus attribute variations in the mode of eruption of this one lava dome, from Vulcanian-type explosion through pyroclastic flows to non-explosive rock fall avalanche, to variable excess pore pressures in the lavas. The extruded lava, although its major element composition has remained nearly constant for more than a year ($\text{SiO}_2 = 65 \text{ wt}\% \pm 0.5 \text{ wt}\%$), may have heterogeneous excess pore pressure due mainly to percolative degassing in the conduit. □

Received 29 June; accepted 27 October 1992.

1. Jaupart, C. & Allegre, C. J. *Earth planet. Sci. Lett.* **102**, 413–429 (1991).
2. Eichelberger, J. C., Carrigan, C. R., Westrich, H. R. & Price, R. H. *Nature* **323**, 598–602 (1986).

3. Rose, W. I., Pearson, T. & Bonis, S. *Bull. Volcanol.* **40**, 23–38 (1977).
4. Fink, J. H. & Manley, C. R. *IAVCEI Proc. Volcanol.* **1**, 169–179 (1989).
5. Fink, J. H. *Nature* **351**, 611 (1991).
6. Ohta, K. *J. geol. Soc. Japan* **97**(7), i–iii (1991).
7. Nakada, S. & Fujii, T. *J. Volcanol. geotherm. Res.* (in the press).
8. Fink, J. H. *Nature* **357**, 119 (1992).
9. Shimozuru, D. *Nature* **353**, 295–296 (1991).
10. Mellors, R. A., Walitt, R. B. & Swanson, D. A. *Bull. Volcanol.* **50**, 14–25 (1988).
11. Self, S., Wilson, L. & Nairn, I. A. *Nature* **277**, 440–443 (1979).
12. Shaw, H. R. *Am. J. Sci.* **272**, 870–893 (1972).
13. Fink, J. H., Anderson, S. W. & Manley, C. R. *J. geophys. Res.* **97**, 9073–9083 (1992).
14. Maeda, S. *et al. Abstr. Fall Mtg. Volcanol. Soc. Japan* **2**, 62 (1991).
15. Wilson, C. J. N. *J. Volcanol. geotherm. Res.* **8**, 231–249 (1980).
16. Sparks, R. S. J. *J. Volcanol. geotherm. Res.* **3**, 1–37 (1978).
17. Seino, M., Fukui, K., Uhira, K., Hashimoto, T. & Yamasato, H. *Abstr. Fall Mtg. Volcanol. Soc. Japan* **2**, 39 (1991).
18. Anderson, S. W. & Fink, J. H. *Nature* **341**, 521–523 (1989).
19. Tomiya, A., Takahashi, E. & Nakada, S. *Abstr. 1992 Japan Earth planet. Sci. Joint Mtg.* 166 (1992).
20. Fink, J. H. *Nature* **352**, 188 (1991).
21. Macdonald, G. A. *Volcanoes* (Prentice-Hall, Englewood Cliffs, New Jersey, 1972).
22. Fisher, R. V. & Schmincke, H.-U. *Pyroclastic Rocks*, 472 (Springer, Berlin, 1984).

ACKNOWLEDGEMENTS. We thank K. Ohta and K. Kamo for their support at Unzen. Comments by D. Swanson, R. S. J. Sparks and J. Fink improved the manuscript. Figure 1 is from a videotape offered by NHK Japan.

Exact tracking of pollen transfer and mating in plants

L. Anders Nilsson*, Elisabeth Rabakonandrianina† & Börge Pettersson*

* Department of Systematic Botany, Villavägen 6, S-752 36 Uppsala, Sweden

† Service de Biologie Végétale et Biochimie, Université d'Antananarivo, BP 906, Antananarivo 101, Madagascar

UNLIKE animals, where individuals engage in direct sexual encounters, higher plants interact sexually only through minute, usually animal-mediated pollen grains, a trait that has hampered understanding of processes that govern plant evolution. A new technique using microtags to mark individual orchid pollinia and monitoring of all stigmas for pollination made it possible to measure exactly pollen transfer and mating pattern in a plant species. We report here that in populations of a hawkmoth-pollinated orchid, *Aerangis ellisii*, pollen transfers were found to be infrequent, to involve single pollen parents, and to occur mostly within 5 metres. Pollinator-mediated patterns of disproportional reproductive success suggest that floral traits are being shaped by mutual sexual selection as proposed by Darwin^{1–3}. The microtag method opens an avenue for novel exploration of plant evolution.

Use of coloured dye particles as pollen analogues has yielded components of fitness for selected plant individuals⁴, whereas for favourable cases genetic paternity-exclusion procedures^{5–7} have allowed pollen parents to be identified for up to 85% of the fruits produced⁸. But the genetic techniques have not permitted the complete mapping and analysis of animal-mediated pollen transfer, effectuated matings and amount of gene flow in whole natural populations. Orchids package their pollen in coherent masses, the pollinia, which are often dispensed intact to stigmas by the pollinator. Orchid pollinia are often solid, 1–2 mm in diameter and only two per flower. Marking pollinia with coloured histochemicals has allowed the study of gene flow between a few individuals within populations⁹. An isolated group of populations of *Aerangis ellisii* (Orchidaceae) in Madagascar was studied throughout its bloom in 1992 (Table 1). A plant produces one to several inflorescences, each with 8–12 white, long-spurred flowers; pollination is by hawkmoths¹⁰. All pollinia in two populations, A and D (containing 148 and 3 flowering individuals, respectively), were marked with microtags. Two control populations, B and C, were also studied.

Monitoring of stigmas revealed the hawkmoth-generated pattern of pollinium transfers (Fig. 1), where a transfer involved one or two donor pollinia being captured by a flower's stigma

(Fig. 2). Transfers were few in all populations (Table 1). In populations A and D, 37% and 67% of the transfers, respectively, involved untagged pollinia representing foreign, imported gene flow. There were no transfers between populations A and D, nor from them to either B or C.

Pollinium dispersal was mostly between adjacent plants (<5 m apart), but 26% travelled >15 m; maximum transfer

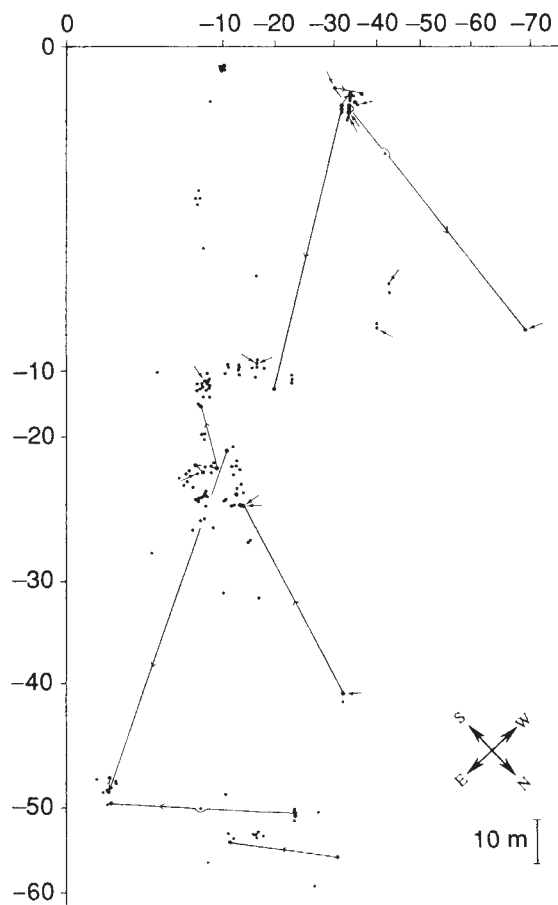


FIG. 1 Pattern of pollinium transfer within the experimental population A of *Aerangis ellisii* growing near the top of Ambatolava rock. Each dot represents an individual plant, and the lines connect plants that participated in tagged transfers. Solitary arrows denote receipt of unmarked pollinia (foreign, imported gene flow). Relative altitude (in m) is indicated along the margins. A few plants (without transfers) that occurred further down the rock have been omitted from the figure.

distance was 76 m (Fig. 3). Most of the transfers were to increasing altitudes on the rock (Fig. 3), probably reflecting the movement of hawkmoths to hilltops at night. About 19% of the transfers were within the same inflorescence, 11% between inflorescences on the same plant, and 70% between different plants (fruit set: 40%, 100% and 84%, respectively). Including foreign transfers, cross-pollination in population A was 83%. Multiple paternity from transfers was absent: 70% involved a single pollinium and 30% two pollinia from a single donor flower. In two (7.7%) of the transfers, the paired pollinia of a donor flower were dispensed onto different stigmas, once within the same and once on different recipient plants. Pollinium transfer from one flower to another almost always (92%) occurred within one night.

Pollinium transfer was not evenly distributed in population A (Fig. 1). For example, one uppermost patch on the rock, representing 11% of the individuals, had 27% of the transfers, the traffic of hawkmoths being apparently concentrated along the forest border. Twenty-nine (19.6%) plants donated or received tagged pollinia; of these, 31% functioned only as males (donation only), 41% only as females (receipt only), and 28% as hermaphrodites (both donation and receipt). Pollinators thus affected the gender of the structurally hermaphroditic plants^{11,12}.

There was a significant relationship between individual donation success and pollinium removal (Spearman rank correlation coefficient $r_s = 0.418$, $P = 0.0001$, $N = 151$). Pollinium removal is therefore a useful measure of male reproductive success in plants, as also indicated by a previous study⁸. Moreover, plants

TABLE 1 Characteristics, and hawkmoth-generated pollinium transfer and fruit-set, of the four populations of *Aerangis ellisii* on the granite outcrop of Ambatolava (1,670 m) in Angavokely, central Madagascar, in 1992

Population	Distance from A (m)	Altitude relative to upper limit of A (m)	No. of flowers produced	No. of transfers involving		Fruit set (%)
				tagged pollinia	untagged pollinia	
A	0	0	1,295	26	15	2.4
B	150	-120	1,195	0	43	1.8
C	800	-160	281	0	22	13.9
D	925	-75	34	1	2	5.9

Population A grew near the top of the rock, B in centre, C far down, and D on an isolated small rock southwest of A. There was no other known occurrence of the plant within 5 km. All pollinia (2,658) in the populations A and D were individually marked with numbered microtags at anthesis. Flowering occurred 16 February–4 April. We checked all open flowers in A and D daily for pollinium transfer, and later for fruit formation; B and C were unmanipulated controls for censusing levels of pollination and fruit set, and were checked every third day. Because a pollinium deposited on a stigma disintegrated and was distinctly visible only during the next 24 h, the numbers of untagged transfers observed in B and C represent minimum estimations. We used a razor blade to open closed stigmas in B and C because presence of microtags would persistently reveal any pollen flow from A and D. Most tagged (78%) as well as untagged (59%) transfers in A and D resulted in fruit, indicating that the microtags did not affect fruit setting.

donating pollen had a higher probability of receiving pollen (being functionally hermaphrodite) than did non-pollen-donating plants ($\chi^2 = 7.56$, $P < 0.01$). Of the total sexual encounters, functionally hermaphrodite plants accounted for 44%, and a few individuals in the population enjoyed relatively high success: 4 plants (2.7%) accounted for 33%, 14 plants (9.5%) for 71% of all sexual encounters. Clearly, hawkmoths made not only the breeding (that is the effective) plant population small, namely 36 individuals (24%), but the reproductive success of individual plants disproportional. Such differential reproductive success arises because of variation in plant phenotypic traits¹³, and may result in pollinator-mediated mate choice between the two sexes in plants¹², as in mutual sexual selection as proposed by Darwin¹⁻³.

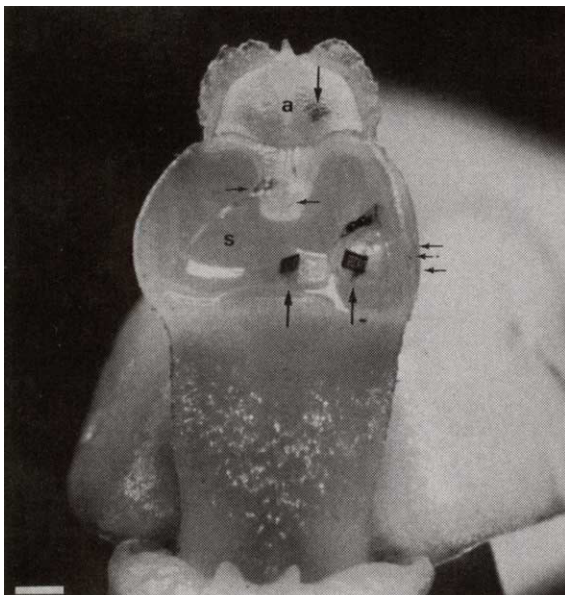


FIG. 2 Individually marked pollinia, transferred by one hawkmoth, in a flower of *Aerangis ellisii*. The flower's large, mucilage-covered stigma (s) has received two pollinia labelled with microtags (upward arrows). Its own pollinia still remain in the anther (a); the tag on the right pollinium is visible through the transparent tissue (downward arrow). The flower parts bear several scales (horizontal arrows) from the hawkmoth pollinator. In addition, a small wasp has perished in the mucilage above the right deposited pollinium. Scale bar, 1 mm.

METHODS. Tags were manually cut out from a microfilm (KODAK Imagelink HQ 3461 SP 615, thickness 0.06 mm) with letter-plus-number combinations (A49, H23 and so on), allowing three characters in breadth within a 0.5 mm tag. Microtags were made self-adhesive by adding minute droplets of glue removed from self-adhesive paper tape (Esselte correction tape) using chloroform and a fine hair-pencil. The protective paper strip on the tape was used for storing ready-to-use microtags. At flower anthesis, we opened the anther cap, placed tags on the two pollinia using an insect pin, and then reclosed the anther. Pollinium size was $\sim 1.1 \times 0.9$ mm.

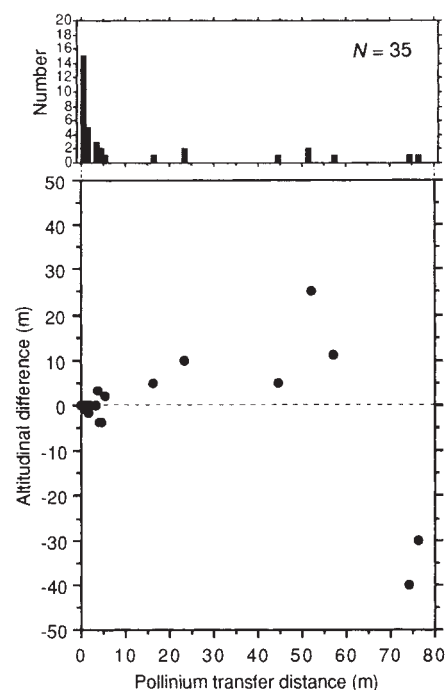


FIG. 3 Distribution of pollinium transfer distances in *Aerangis ellisii* as revealed by individually tagged pollinia. Transfers mostly occurred over short distances, with the distribution being leptokurtic (top), and in a direction of increasing altitude on the rock (bottom).

The use of microtags opens up a new avenue for obtaining precise information on animal-mediated sexual interactions in plants, and offers the further possibility of measuring lifetime mating success. Our method emphasizes the usefulness of pollinium-bearing plants such as orchids as model organisms for investigating plant evolutionary processes, as Darwin suggested 130 years ago¹⁴. □

Received 8 June; accepted 23 October 1992.

1. Darwin, C. *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).
2. Huxley, J. S. *Proc. zool. Soc. Lond.* **35**, 491–562 (1914).
3. Arnold, S. J. In *The Evolution of Sex and its Consequences* (ed. Stearns, S. C.) 283–315 (Birkhäuser, Basel, 1987).
4. Campbell, D. R. *Evolution* **43**, 318–334 (1989).
5. Ellstrand, N. C. *Am. Nat.* **123**, 819–828 (1984).
6. Meagher, T. R. *Am. Nat.* **128**, 199–215 (1986).
7. Devlin, B. & Ellstrand, N. C. *Evolution* **44**, 248–259 (1990).
8. Broyles, S. B. & Wyatt, R. *Evolution* **44**, 1454–1468 (1990).
9. Peakall, R. & Beattie, A. J. *Evolution* **45**, 1837–1848 (1991).
10. Nilsson, L. A. & Rabakonandrianina, E. *Bot. J. Linn. Soc.* **97**, 49–61 (1988).
11. Lloyd, D. G. & Bawa, K. S. *Evol. Biol.* **17**, 255–339 (1984).
12. Nilsson, L. A. *Evol. Trends Pl.* **6**, 33–40 (1992).
13. Nilsson, L. A. *Nature* **334**, 147–149 (1988).
14. Darwin, C. *On the Various Contrivances by which British and Foreign Orchids are Fertilised by Insects* (Murray, London, 1862).

ACKNOWLEDGEMENTS. We thank Malagasy students for most of the marking/monitoring on the steep rock: E. O. Rasoaandrasana, N. F. Rakotomanga, E. Razafimandrisoa, R. Randrianaivo, E. Rakotonandrianina, H. Ralamaro, A. Rakotondrafara, J. Rakotoarimanana and R. Randriambololona. H. Dobson commented on the manuscript. The Swedish Natural Research Council (NFR) and the Swedish Agency for Research Cooperation with Developing Countries (SAREC) provided support.

Rapid microevolution of migratory behaviour in a wild bird species

P. Berthold*, A. J. Helbig*†, G. Mohr* & U. Querner*

* MPI Verhaltensphysiologie, Vogelwarte Radolfzell, D-7760 Möggingen, Germany

THE Blackcap, *Sylvia atricapilla*, a widespread Palearctic migratory bird, rarely wintered in Britain until the 1950s. The winter population has since increased to several thousand birds^{1,2}. Ringing indicates that these are not British Blackcaps forestalling migration, but birds breeding in Continental Europe reaching Britain on a novel westerly migration route^{3,4}. The proportion of north-western migrants among Blackcaps ringed in parts of Germany and Austria has increased from 0% before 1960 to currently 7–11%^{5–7}. We bred British wintering Blackcaps in captivity and determined the migratory direction of their offspring. Here we report that these birds migrate west-northwest in autumn, a direction genetically distinct from the British breeding population and the predominantly southwestern migratory population of west-central Europe. The novel route must have evolved within the past 30 years with selection favouring birds wintering some 1,500 km further north than most of their conspecifics. To our knowledge, this is the first case in any vertebrate in which a drastic and recent evolutionary change of behaviour has been documented and its genetic basis established.

To investigate whether the novel migration route has a genetic basis, we caught 20 Blackcaps of each sex near Weston-super-Mare (51°17'N, 3°00'W; Sedgemoor District, England) in December and January in 1988–1990 and transferred them to Radolfzell (southwest Germany). They were kept indoors until early March, when pairs were released into outdoor aviaries. In two breeding seasons seven pairs produced 41 young, which were hand-raised from the age of 6–8 days, as was a control group of 49 nestlings taken from nests near Radolfzell. Each individual's migratory orientation was tested 15–20 times (20

September–17 November) outdoors at dusk at the rearing site using a modified Emlen technique^{8,9}.

Adult birds that wintered in Britain and their F₁ offspring oriented slightly north of west (means 279° and 273°; Fig. 1), and are statistically indistinguishable with respect to mean direction and dispersion of individual vectors (Table 1). Both differed from southwest German controls, which oriented southwest (227°), corresponding to the classical migration route from southwest Germany to the western Mediterranean region (compare with the compass course for Radolfzell–Sevilla, Spain; Fig. 2). The same conclusions hold if overall mean directions are calculated for the offspring of each pair, assuming that individual vectors of siblings may not be quite independent. Offspring of pairs wintering in Britain ($N = 7$; 275°) differed from groups of nest mates from Germany ($N = 18$; 230°; $P < 0.001$; Watson–Williams test). As a group, adults with migratory experience showed more concentrated orientation than juveniles, although the difference was not significant (Table 1). Among captive-bred juveniles, most variation of orientation was between rather than within sibling groups. Although the circular distribution of the data precludes an ANOVA, the 9 significant differences of mean direction between non-siblings (7 cases at $P < 0.05$, 2 at $P < 0.01$), but none between siblings, indicate a genetic component to the variance in orientation.

Orientation of hand-raised passerines in orientation funnels usually reflects genetically programmed migratory directions^{6,10}. Our results show that Blackcaps wintering in Britain differ in innate migratory direction from (1) the southwest-migrating

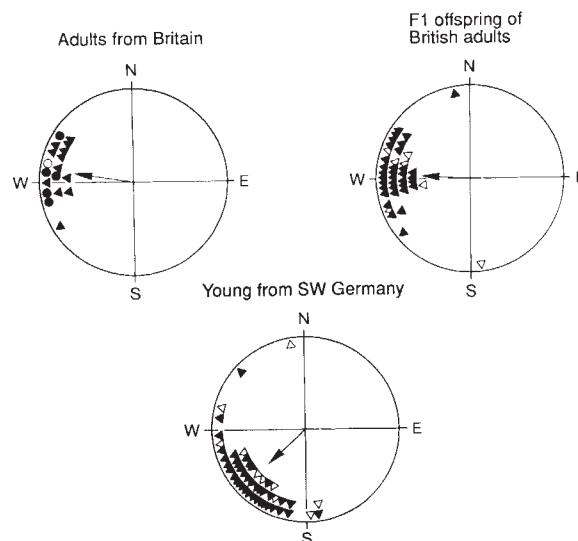


FIG. 1. Orientation of Blackcaps caught in winter in Britain and tested the following autumn ('adults'), their F₁ offspring (F₁) and a control group of hand-raised birds from Radolfzell (47°46'N, 8°59'E), Germany. Each symbol gives the mean direction of one bird during 15–20 tests (filled symbol, vector significant at $P < 0.05$, Rayleigh test; open symbol $P > 0.05$). Among adults triangles show birds that are parents of F₂ offspring in upper right diagram (two adults were involved in two pairings each), dots show birds with no reproductive success. Arrows, group mean vectors.

METHODS. Before orientation-testing, young birds were individually placed into wire cages in outdoor aviaries at the prospective test site for at least 8 weeks. Here they experienced the natural photoperiod and day and night sky and got used to some distant artificial light sources. Orientation tests: Each bird was placed in an aluminium funnel (top diameter 35 cm) lined with typewriter correction paper and covered with clear glass (no peripheral shields). Starting at local sunset, the orientation of individuals from all 3 groups was tested simultaneously for 1.5–2 h to ensure that potential external influences did not differentially affect them. Data analysis: Scratches left by the birds on funnel papers were counted in 24 sectors on a light screen. For each active test (40+ scratches) vector addition yielded a directional heading, from which mean vectors of orientation were calculated for each individual across the entire season. Statistical analysis follows ref. 18. For treatment of axially bimodal activity distributions, see ref. 9.

† To whom correspondence should be addressed at: Institut für Pharmazeutische Biologie, Universität Heidelberg, Im Neuenheimer Feld 364, D-6900 Heidelberg, Germany.

majority of the Central European breeding population, (2) Blackcaps breeding in Briatin, not tested here, but known from ringing to migrate exclusively southward¹¹ and (3) from south-west-migrating ($\sim 222^\circ$, Fig. 2) Scandinavian Blackcaps, which pass through Britain in autumn, but have not yielded any winter recoveries there. Note that this conclusion is based entirely on the behaviour of captive-bred juveniles. Orientation in captivity of experienced adults could be interpreted as homing towards familiar winter quarters, a capacity demonstrated by displacement experiments with Starlings¹². But the concordance of mean orientation between adult experimental Blackcaps and their offspring shows that a genetically programmed westerly direction originally took these birds to Britain and was inherited in the F_1 generation. Overall mean directions suggest a breeding origin between Belgium and central Germany, where many west-northwest-migrants have indeed been ringed^{5,6}.

Alternative hypotheses can be excluded as follows: first, Blackcap numbers caught at British bird observatories in autumn (partly of Scandinavian origin) correlate with those wintering subsequently¹, suggesting that some winterers may be from Scandinavia. But no ringing recoveries support this and the correlation could instead result from parallel year-to-year variation of breeding success between northern and central Europe, thus not supporting a Scandinavian origin. Second, some Scandinavian or continental southwest-migrants could regularly drift off course in southeasterly winds and overwinter in Britain because of improved conditions, thus reaching their destination because of environmental rather than genetic influences. Both hypotheses can be rejected on the basis of our breeding experiment, because in either case a more southerly innate migratory direction ($\sim 225^\circ$, Fig. 2) would be expected for the F_1 offspring of Blackcaps wintering in Britain. Finally, one may accept their

TABLE 1 Second-order mean vectors of orientation of the three groups of Blackcaps tested

Group	N	r	a
1 British wintering adults*	18	0.65	279°
2 F_1 offspring of British adults*	41	0.53	273°
3 Young from SW-Germany nests†	49	0.56	227°

The orientation of each group differs significantly from random ($P < 0.001$, Hotelling one-sample test). N, number of birds; r, vector length; a, mean direction.

* Groups 1 and 2 differ neither in mean direction ($P > 0.05$, Watson-Williams test) nor scatter ($P > 0.05$; Mann-Whitney U test).

† Mean direction of group 3 differs from 1 and 2 at $P < 0.001$ (Watson-Williams test).

continental origin, but assume that climate amelioration¹³ led to relaxed selection on genetically determined migration directions, increasing the portion of birds choosing opportunistically where to migrate. Choices may depend on prevailing winds or on resource distribution, which could be assessed by exploratory flights with little genetically programmed directionality (compare with the 'familiar area' concept of Baker¹⁴). But offspring from British wintering Blackcaps oriented as consistently in funnel cages as German controls (Fig. 1), and individual mean vectors were only slightly shorter in 'British-bred' versus 'German' juveniles (median $r = 0.604$ versus 0.655 ; $P < 0.05$), a small difference, unlikely to indicate less strict genetic control.

We conclude that during the past three decades some central European Blackcaps have evolved a novel migration route establishing new winter quarters 1,000–1,500 km north of the traditional west Mediterranean wintering areas. This shift must be due to an increased frequency of distinct genotypes, because northwest-migrants were not noted until 1961⁴ (although thousands had been ringed before in central Europe), but now account for 7–11% of the breeding population in parts of Germany and Austria^{5–7}. A high evolutionary potential for other migratory adaptations in Blackcaps had previously been demonstrated by selective breeding¹⁵. West to northwest migratory directions may fall within the normal range of genetic variation and probably occurred at a low frequency in the past. The disproportional increase of west-northwest-migrants, however, is probably due to selection by improved wintering conditions that led to lowered winter mortality in Britain. Potential advantages of the new winter quarters include improved winter food (for example, bird tables that are used intensively), climate amelioration, shorter migration distance, less intraspecific competition, earlier return to breeding areas due to an advanced spring photoperiod, and physiological preadaptation to potentially harsh conditions on the breeding grounds in early spring¹⁶. Photoperiodic simulations suggested that arrival on the breeding grounds from Britain should be about 10 days earlier than from Mediterranean latitudes¹⁷. Assortative mating among early arrivals may be an important mechanism driving the disproportional increase of west-northwest-migrants. Year-round residency has not yet evolved in British breeding Blackcaps, but this may be only a matter of time. A drastic shortening of endogenously programmed migratory distance may require more complex physiological adaptations and evolve more slowly than changes of migratory direction. □



FIG. 2. Map of northwest-Europe showing potential migration routes of Blackcaps breeding, migrating through or wintering in Britain. W, Weston-super-Mare (origin of test birds); R, Radolfzell (origin of control group, test site); Bo, Bonn; Ba, Bayreuth; O, Oslo.

Received 17 June; accepted 19 October 1992.

1. Leach, I. H. *Bird Study* **28**, 5–14 (1981).
2. Lack, D. *The Atlas of Wintering Birds in Britain and Ireland* (Poyser, Calton 1986).
3. Langslow, D. R. *Bird Study* **26**, 239–252 (1979).
4. Zink, G. *Vogelwarte* **21**, 222–223 (1962).
5. Berthold, P., Querner, U. & Schlenker, R. *Die Mönchsgrasmücke* (Ziensen, Wittenberg, 1990).
6. Helbig, A. J. *Oecologia* **90**, 483–488 (1992).
7. Helbig, A. J. *J. evol. Biol.* **14**, 657–670 (1991).
8. Emlen, S. T. & Emlen, J. T. *Auk* **83**, 361–367 (1966).
9. Helbig, A. J. in *Orientation in Birds* (ed. Berthold, P.) 270–306 (Birkhäuser, Basel, 1991).
10. Wiltschko, W. & Gwinner, E. *Naturwissenschaften* **61**, 406 (1974).

11. Zink, G. *Der Zug europäischer Singvögel. Ein Atlas der Wiederfunde beringter Vögel* Vol. 1 (Vogelzug, Möggingen, 1973).
12. Perdeck, A. C. *Ardea* **46**, 1–37 (1958).
13. Rocznik, K. *Naturwiss. Rundschau* **42**, 361–363 (1989).
14. Baker, R. R. *The Evolutionary Ecology of Animal Migration* (Hodder & Stoughton, London, 1978).
15. Berthold, P., Mohr, G. & Querner, U. *J. Ornith.* **131**, 33–45 (1990).
16. Berthold, P. & Terrill, S. B. *Ring. Migr.* **9**, 153–159 (1988).
17. Terrill, S. B. & Berthold, P. *Oecologia* **85**, 266–270 (1990).
18. Batschelet, E. *Circular Statistics in Biology* (Academic, London, 1981).

ACKNOWLEDGEMENTS. We thank the British Trust for Ornithology and the Nature Conservancy Council for their cooperation and permission to catch Blackcaps in Britain. The help of C. Mead, G. Pudney, T. Parsons, comments by H. Dingle and financial support by Deutsche Forschungsgemeinschaft are acknowledged.

Atherogenesis in transgenic mice expressing human apolipoprotein(a)

Richard M. Lawn, David P. Wade, Robert E. Hammer*, Giulia Chiesa†, Judy G. Verstuyft‡ & Edward M. Rubin‡

Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, California 94305, USA

* Howard Hughes Medical Institute, † Department of Molecular Genetics, University of Texas Southwestern Medical School, Dallas, Texas 75235, USA

‡ Division of Cell and Molecular Biology, Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720, USA

ELEVATED plasma levels of the lipoprotein Lp(a) are associated with increased risk for atherosclerosis and its manifestations, myocardial infarction, stroke and restenosis (for reviews, see refs 1–3). Lp(a) differs from low-density lipoprotein by the addition of the glycoprotein apolipoprotein(a), a homologue of plasminogen that contains many tandemly repeated units which resemble the fourth kringle domain of plasminogen, and single homologues of its kringle-5 and protease domain⁴. As plasma Lp(a) concentration is strongly influenced by heritable factors and is refractory to most drug and dietary manipulation, the effects of modulating it are difficult to mimic experimentally. In addition, the absence of apolipoprotein(a) from virtually all species other than primates precludes the use of convenient animal models. Here we show that transgenic mice expressing human apolipoprotein(a) are more susceptible than control mice to the development of lipid-staining lesions in the aorta, and that apolipoprotein(a) co-localizes with lipid deposition in the artery walls.

A quantitative model has been developed for scoring lipid-staining regions of the mouse aorta that resemble the 'fatty streaks' thought to represent early atherosclerotic lesions in humans^{5–9}. Certain inbred strains of mice, such as C57BL/6, are especially susceptible to the development of these lipid lesions after only 8–14 weeks on a high-cholesterol diet, whereas most inbred or hybrid mice are resistant⁵. Both coronary arteries and the proximal ascending aorta of susceptible strains of mice consistently develop lipid-staining lesions when fed atherogenic diets, but because of its larger size, the aorta is more conveniently examined for quantitative studies. The areas of fatty-streak lesions occurring in the first three months of fat feeding are usually assessed, although susceptible strains of mice retained on high-cholesterol diets for 8–9 months do go on to develop fibro-fatty lesions which have many of the characteristics of advanced human atheromatous plaques^{5,10}.

Mice transgenic for human apolipoprotein(a) (referred to as apo(a)) were made by introducing apo(a) complementary DNA linked to the mouse transferrin promoter into a hybrid genetic background of strains C57BL/6 crossed with SJL¹¹. The transgene is expressed primarily in liver, and to a lesser degree in other tissues¹¹. It results in a murine plasma apo(a) concentration of ~10 mg dl⁻¹, which is comparable to the median level in humans. To determine whether the expression of human apo(a) could alter the atherogenic susceptibility of mice, transgenic and control littermates were fed an atherogenic diet

containing 1.25% cholesterol and 7.5% saturated fat¹² for about 3.5 months. Animals were then killed and aortic sections prepared for analysis as in ref. 7, modified according to ref. 9. Sectioning and lesion scoring were done without knowledge of the transgenic or control status of the sample. The mean lipid-staining lesion area per mouse was 1,072 μm^2 for transgenic mice and 58 μm^2 for control littermates. The groups are statistically different at $P < 0.001$ by Mann-Whitney U -test analysis. The effect was dependent on the atherogenic diet, as transgenic mice maintained on the low-fat diet had a mean lesion area of only 19 μm^2 (Fig. 1).

On the atherogenic diet, 14 of the 15 transgenic mice developed lesions in scored sections, and the single transgenic mouse with a zero lesion area measurement did contain a large lesion which was just proximal to the section where scoring began. Of the 60 sections scored for the transgenic group (4 per animal), 38 contained one or more lesions. Of the 8 non-transgenic animals, 6 were lesion-free and 2 contained single small aortic lesions (2 of 32 scored sections contained single lesions). In addition to measuring lesion area per section, we noted the number of separate lesions in each section examined. The mean number of lesions per section was 1.25 for the transgenic and 0.06 for the control groups. The variation in lesion score among

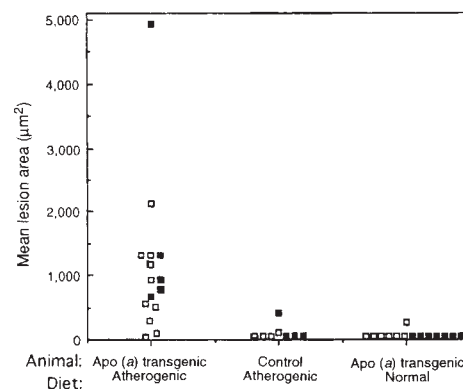


FIG. 1 Lipid-staining lesion area in mouse aortas. The mean lesion area per section is shown for each mouse on the atherogenic or normal diets. Males are indicated by filled squares and females by open squares. The median lesion area for the group of transgenic mice on the atherogenic diet was 1,072 μm^2 and for the non-transgenic group on the atherogenic diet, 58 μm^2 ; the median values were 897 and 0, respectively. These groups are statistically different by the non-parametric Mann-Whitney U -test at $P < 0.001$. Mean plasma concentration of apo(a) in the transgenic mice was 10 (± 2.6) mg dl⁻¹ ($N = 15$) and zero in the non-transgenics ($N = 8$); total cholesterol was 154 (± 10.3) mg dl⁻¹ in the transgenic and 175 (± 5.3) mg dl⁻¹ in the non-transgenic mice on the atherogenic diet. This atherogenic diet does not affect apo(a) levels, but, as shown in ref. 12, it raises total cholesterol by ~70 mg dl⁻¹, with the increase occurring in the density ranges of VLDL to LDL.

METHODS. Transgenic mice were produced by expressing human apo(a) cDNA (isoform $M_r = 550,000$; ref. 30) with the mouse transferrin promoter as described¹¹. The founder transgenic mouse derived from C57BL/6 $\times$ SJL F_2 eggs was mated to a C57BL/6 $\times$ SJL F_1 to establish a line. The line was maintained by continued crossing to C57BL/6 $\times$ SJL hybrids. Control mice were non-transgenic littermates. After 6 weeks of age, transgenic mice and non-transgenic siblings were switched from a normal mouse chow diet (Techlad, Madison; 4% mouse/rat chow) to an atherogenic diet containing 1.25% cholesterol, 7.5% saturated fat as cocoa butter, 7.5% casein and 0.5% sodium cholate¹² for ~3.5 months. At that time, plasma apo(a) was determined by ELISA (GeneScreen) and total cholesterol by the cholesterol oxidase method¹². For each animal, four 10- μm sections separated by 50 μm were quantified^{7,9}. The first and most proximal section to the heart was taken 80 μm distal to the point where the aorta becomes rounded. The area of Oil Red O-staining in each section was measured with a calibrated eyepiece, and the mean lesion area per section per animal was calculated for each individual and group of animals. Coded slides were examined blind in two separate analyses by the same examiner and gave consistent results (correlation coefficient $R > 0.92$).

transgenic animals showed no obvious correlation with the plasma concentration of apo(a) or cholesterol in the mice. The plasma concentration of high-density lipoprotein (HDL) did not differ significantly between the transgenic ($53.8 \pm 5.0 \text{ mg dl}^{-1}$) and non-transgenic ($44.6 \pm 3.3 \text{ mg dl}^{-1}$) littermates on the high-fat diet.

In humans, apo(a) has been identified within atherosclerotic plaque, but not in adjacent 'unaffected' regions of the vessel wall¹³⁻¹⁷. To see if apo(a) was associated with the lipid-staining lesions in the transgenic mice, sections were immunostained with an anti-human apo(a) antibody. This antibody did not crossreact with mouse plasminogen on a western blot or on immunostaining of livers of normal and transgenic mice (data not shown). Apo(a) immunoreactivity was apparent in localized regions of the aorta, which in each case corresponded to regions of lipid staining in adjacent or closest available sections. Detectable apo(a) immunoreactivity was not an overall feature of the vessels, but was concentrated at the lesion site (Fig. 2). These observations support the role of apo(a) in lesion development and argue against the possibility that the apo(a) transgene is generally active in aortic cells. The regions of the aorta that stain for neutral lipids and apo(a) antigen appear intimal and medial in nature. Higher magnification reveals the presence of

macrophage-like cells, a common feature of fatty streaks in humans.

In human plaques, apolipoprotein B-100 (apoB-100) is associated with apo(a). ApoB-100 is a component of both Lp(a) and non-apo(a)-containing lipoproteins such as low- and very-low-density lipoprotein (LDL and VLDL) particles. ApoB-100 can become associated with atherosclerotic plaques in its own right, but apo(a) is more highly concentrated in vessels than it is in plasma, compared with apoB-100 (refs 13-17). This implies that apo(a) may have preferential binding sites in the vessel wall, and that much of the apoB-100 in plaques could be due to its association with apo(a). We detected apoB-immunostaining in lipid- and apo(a)-staining regions of transgenic aortas using anti-(rat)apoB antibody (Fig. 3). In humans, virtually all of the apo(a) in plasma is present in the form of lipoprotein particles. To determine whether apo(a) was lipid-associated in these animals, plasma from three transgenic mice on the atherogenic diet was centrifuged at a density of 1.215 g ml^{-1} followed by western blotting with anti-apo(a) antibody. Only 5% of the total apo(a) from mouse plasma was detected in the lipid-floating fraction, with the majority pelleting in a lipid-free form (Fig. 4). These results are consistent with other studies¹¹ that demonstrate weak association of transgenic apo(a) with mouse LDL compared to human LDL.

In our study, the apo(a) transgenic and non-transgenic control littermates were of similar, but not identical, genetic background. Earlier results from inbred strains have concluded that reduced HDL levels are an invariant marker for atherosclerosis

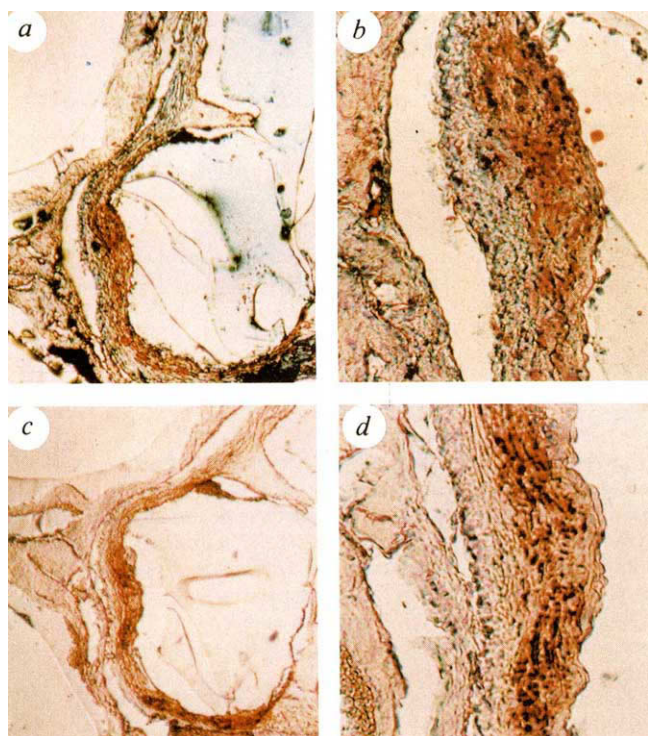


FIG. 2. Lipid and apo(a) accumulation in the aorta of transgenic mice. Low- (50 $\times$) (a, c) and high- (150 $\times$) magnification views of adjacent sections of proximal aorta from an apo(a) transgenic mouse fed an atherogenic diet were stained for lipid (in a and b) and apo(a) (in c and d).

METHODS. The heart and attached aorta were fixed in 10% formalin and 10- μm sections were prepared⁷. Lipid sections were stained with Oil Red O and haematoxylin, and counterstained with light green⁸. For apo(a) immunostaining, sections were pre-incubated with 1% rabbit serum for 1 h, and subsequently with 1:1,000 sheep anti-human Lp(a) antibody absorbed against plasminogen (Immuno, Vienna). This antibody has no appreciable cross-reactivity against mouse plasminogen as determined by western blotting and by immunostaining of control and transgenic mouse liver sections. Bound primary antibody was detected by incubation with biotinylated rabbit anti-sheep IgG at 1:200 followed by quenching of endogenous peroxidase activity with 0.6% H_2O_2 , incubation with avidin-biotinylated peroxidase complex, and colour production with 3-amino-9-ethyl carbazole following the vendors protocols (Vector Lab). Control slides were prepared with non-immune sheep serum in place of the primary antibody.

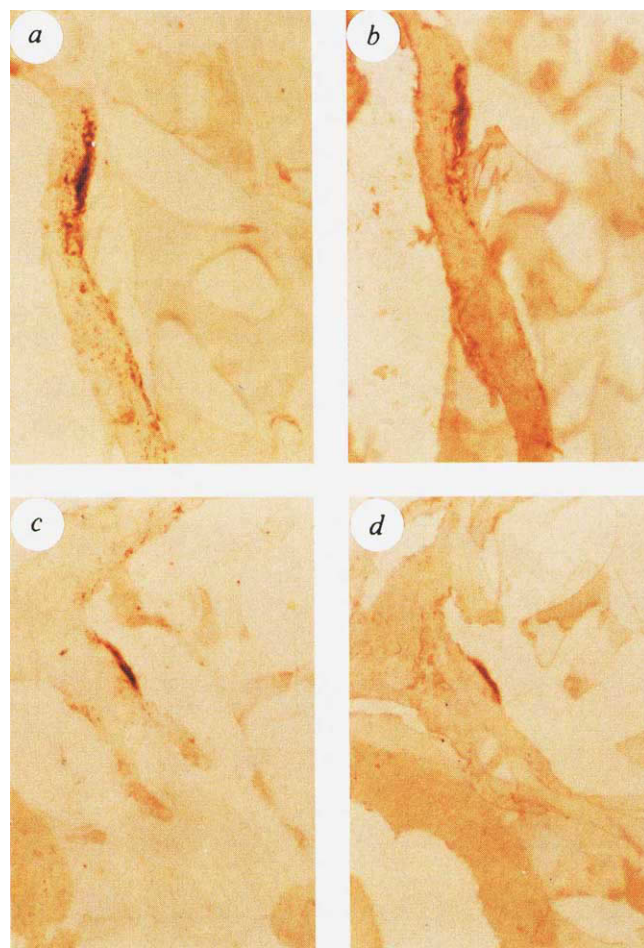


FIG. 3. Co-localized apoB and apo(a) immunostaining of transgenic aorta. Two examples are shown of adjacent sections of aorta stained for apo(a) (in parts a and c) or apoB (in b and d) at 100 $\times$ magnification. The protocol in Fig. 2 legend was followed for apo(a) and modified for use of rabbit primary antibody at 1:500 dilution for apoB.



FIG. 4. Immunoblot of free and lipid-associated apo(a) from the plasma of transgenic mice on the atherogenic diet. Lane P: Total plasma (10 μ l); lanes 1–5, lipid-free (density > 1.215 g ml^{-1}) fraction, adjusted to final volume of 500 μ l; lane 1, 10 μ l; lane 2, 7.5 μ l; lane 3, 5 μ l; lane 4, 2.5 μ l; lane 5, 1.25 μ l. Lane 6 contains 75 μ l out of 500 μ l of the lipoprotein fraction of density < 1.215 g ml^{-1} .

METHODS. Pooled plasma (500 μ l) from three apo(a) transgenic mice on the atherogenic diet was adjusted to a density of 1.215 g ml^{-1} with KBr and centrifuged. After adjusting to equal volumes, aliquots of the top and bottom fractions were electrophoresed on a 6% SDS-polyacrylamide gel under reducing conditions, transferred to nitrocellulose and probed with sheep anti-human apo(a) (Immuno, Vienna) at 1:400, followed with peroxidase-conjugated anti-goat IgG at 1:400, and colour development with 4-chloro-1-naphthol. The ratio of apo(a) in dilutions of the bottom (lipid-free) and top fractions was used to estimate that $\sim 5\%$ of the plasma apo(a) is lipid-associated.

susceptibility in mice⁶. This would not account for the differences seen here, in which control and transgenic groups have similar HDL levels. The highly significant differences in lesion areas between the groups, and the co-localization of apo(a) to the lesions in the transgenic animals, strongly indicate that the difference in atherogenesis between these two groups is due to apo(a). We cannot be certain how the expression of the apo(a) transgene triggers fatty streak formation. The significant 'atherogenesis' in the transgenic mice, despite the fact that only 5% of the plasma apo(a) is found in the lipid float, suggests that apo(a) could effect vessel pathophysiology independently of the apo B-100 and lipid portions of the Lp(a) lipoprotein. It is possible that the fraction of the transgenic apo(a) that is lipid-associated binds to sites in the vessel in the same way as it would in humans. Alternatively, free apo(a) might bind to components in the vessel wall for which it has affinity, such as fibrin, fibronectin, collagen, elastin, glycosaminoglycans, endothelial cells or macrophages^{18–24}, and this immobilized apo(a) might then enhance the deposition of lipoproteins by virtue of the high affinity of apo(a) for apoB-100-containing lipoproteins²⁵. These results also indicate that the relationship of Lp(a) to atherosclerosis may result in part from binding of free apo(a) to sites in the vessel wall followed by antifibrinolytic^{19,20,26–28}, pro-cell migratory²⁹ or inflammatory activities, which could ultimately lead to hyperplasia and macrophage recruitment and the appearance of lipid-containing lesions. Further studies with transgenic animals expressing human apo(a) as well as other apolipoprotein genes are necessary to sort out these possibilities. □

Received 7 September; accepted 13 October 1992.

- Utermann, G. *Science* **246**, 904–910 (1989).
- Scanu, A. M. & Fless, G. M. *J. clin. Invest.* **85**, 1709–1715 (1990).
- Scanu, A. M., Lawn, R. M. & Berg, K. *Ann. Intern. Med.* **115**, 209–218 (1991).
- McLean, J. W. *et al. Nature* **330**, 132–137 (1987).
- Paigen, B., Ishida, B., Verstuyft, J., Winters, R. & Albee, D. A. *Arteriosclerosis* **10**, 316–323 (1990).
- Paigen, B., Mitchell, D., Reue, K., Morrow, A. & Le Bouef, R. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3763–3767 (1987).
- Paigen, B., Morrow, A., Holmes, P., Mitchell, D. & Williams, R. *Atherosclerosis* **68**, 231–240 (1987).
- Mehrebian, M., Demer, L. & Lusis, A. *Arterio. Thromb.* **11**, 947–954 (1991).
- Rubin, E., Krauss, R., Spangler, E., Verstuyft, J. & Clift, S. *Nature* **353**, 265–267 (1991).
- Stewart-Phillips, J. L. & Lough, J. *Atherosclerosis* **90**, 211–218 (1991).
- Chiesa, G. *et al. J. biol. Chem.* (in press).
- Yokode, M., Hammer, R., Ishibashi, S., Brown, M. & Goldstein, J. *Science* **250**, 1273–1275 (1990).
- Rath, M. *et al. Arteriosclerosis* **9**, 579–592 (1989).
- Cushing, G. L. *et al. Arteriosclerosis* **9**, 593–603 (1989).

- Niendorf, A. *et al. Virchows Archiv. A. Path. Anat.* **417**, 105–111 (1990).
- Smith, E. B. *Atherosclerosis* **84**, 173–181 (1990).
- Pepin, J. M., O'Neil, J. A. & Hoff, H. F. *J. Lipid Res.* **32**, 317–327 (1991).
- Harpel, P. C., Gordon, B. R. & Parker, T. S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 3847–3851 (1989).
- Hajjar, K. A., Gavish, D., Breslow, J. L. & Nachman, R. L. *Nature* **339**, 303–305 (1989).
- Miles, L. A., Fless, G. M., Levin, E. G., Scanu, A. M. & Plow, E. F. *Nature* **339**, 301–303 (1989).
- Salonen, E., Jauhainen, L., Zardi, A., Vaheri, A. & Enholm, C. *EMBO J.* **8**, 4035–4040 (1989).
- Zioncheck, T. F., Powell, L. M., Rice, G. C., Eaton, D. L. & Lawn, R. M. *J. clin. Invest.* **87**, 767–771 (1991).
- McConathy, W. J. *et al. FASEB J.* **4**, A1019 (1990).
- Bihari-Varga, M., Gruber, E., Rotheneder, M., Zechner, R. & Kostner, G. M. *Arteriosclerosis* **8**, 851–857 (1988).
- Trieu, V. N., Zioncheck, T. F., Lawn, R. M. & McConathy, W. J. *J. biol. Chem.* **9**, 5480–5485 (1991).
- Edelberg, J. M., Gonzalez-Gronow, M. & Pizzoo, S. V. *Thromb. Res.* **57**, 155–162 (1990).
- Loscalzo, J., Weinfeld, M., Fless, G. M. & Scanu, A. M. *Arteriosclerosis* **10**, 240–245 (1990).
- Rouy, D., Grailha, P., Nigon, F., Chapman, J. & Angles-Cano, E. *Arteriosclerosis Thromb.* **11**, 629–638 (1991).
- Kojima, S., Harpel, P. C. & Rifkin, D. B. *J. Cell Biol.* **113**, 1439–1445 (1991).
- Koschinsky, M. L. *et al. Biochemistry* **30**, 5044–5051 (1991).

ACKNOWLEDGEMENTS. We thank R. Davis and J. Drisko for apo B antibody, and H. Hobbs and T. Redgrave for discussion. Animals used in this study were cared for in accordance with Institutional guidelines. This research was supported by grants from the Cigarette and Tobacco Surtax Fund of the State of California to R.M.L.; from the National Dairy Promotion Research Board to E.M.R.; from the Howard Hughes Medical Institute to R.E.H.; and the Italian Ministry of Scientific and Technical Research to G.C.

Mutation of the β -amyloid precursor protein in familial Alzheimer's disease increases β -protein production

Martin Citron[†], Tilman Oltersdorf^{†*}, Christian Haass, Lisa McConlogue^{*}, Albert Y. Hung, Peter Seubert^{*}, Carmen Vigo-Pelfrey^{*}, Ivan Lieberburg^{*} & Dennis J. Selkoe

Department of Neurology and Program in Neuroscience, Harvard Medical School, and Center for Neurologic Diseases, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA
* Athena Neurosciences Inc., South San Francisco, California 94080, USA

PROGRESSIVE cerebral deposition of the 39–43-amino-acid amyloid β -protein ($A\beta$) is an invariant feature of Alzheimer's disease which precedes symptoms of dementia by years or decades. The only specific molecular defects that cause Alzheimer's disease which have been identified so far are missense mutations in the gene encoding the β -amyloid precursor protein (β -APP) in certain families with an autosomal dominant form of the disease (familial Alzheimer's disease, or FAD)^{1–5}. These mutations are located within or immediately flanking the $A\beta$ region of β -APP, but the mechanism by which they cause the pathological phenotype of early and accelerated $A\beta$ deposition is unknown. Here we report that cultured cells which express a β -APP complementary DNA bearing a double mutation (Lys to Asn at residue 595 plus Met to Leu at position 596) found in a Swedish FAD family⁵ produce ~ 6 – 8 -fold more $A\beta$ than cells expressing normal β -APP. The Met 596 to Leu mutation is principally responsible for the increase. These data establish a direct link between a FAD genotype and the clinicopathological phenotype. Further, they confirm the relevance of the continuous $A\beta$ production by cultured cells^{6–8} for elucidating the fundamental mechanism of Alzheimer's disease.

Human kidney 293 cells were transiently transfected with DNA constructs encoding either wild-type β -APP₆₉₅ (ref. 9) or β -APP₆₉₅ containing the Swedish FAD double mutation⁵, in which the two amino acids immediately amino-terminal to the $A\beta$ region (595 and 596 in β -APP₆₉₅, or 670 and 671 in β -APP₇₇₀) are changed (Fig. 1a). Subconfluent cultures of each transfected cell type were metabolically labelled with [³⁵S]methionine for 16 hours and the conditioned medium immunoprecipitated with a high-titre antibody (R1280) against synthetic $A\beta_{1-40}$ (ref. 6). Fluorography of the precipitates consistently showed substantially more $A\beta$ in the medium from

[†]M.C. and T.O. made equal contributions to this work.

cells expressing the mutation (Fig. 1b). The A β peptide produced by both cell types had all of the characteristics recently described for *in vitro*-generated A β (ref. 6), including co-migration with synthetic A β_{1-40} (Fig. 1b), greater production by β -APP-transfected than by mock-transfected cells (Fig. 2), and failure to immunoprecipitate with the amino-terminal β -APP antibody, B5, and the carboxy-terminal antibody, C7 (data not shown). The increase in A β production in the Swedish transfectants was accompanied by a decrease in the peptide of M_r 3,000 (3K) (known as p3), which starts at residue 17 or 18 of A β (ref. 6) and represents a fragment of the 10K carboxy-terminal peptide that arises from constitutive secretory cleavage of β -APP¹⁰⁻¹⁴.

The increase in A β in the medium from the Swedish transfectants was quantitated using an A β -specific sandwich enzyme-linked immunosorbent assay (ELISA)⁷. The levels of A β detected by this assay were normalized to those of secreted soluble β -APP (APP_s¹¹⁻¹⁴) detected in the same medium by a sandwich ELISA specific for APP_s (Fig. 2). Cells expressing the mutant β -APP₆₉₅ construct produced 6-7-fold more A β than identically transfected cells expressing wild-type β -APP₆₉₅ (Fig. 2). This marked increase was consistently obtained in all transfections with the mutant construct. Moreover, a 7-8-fold increase of A β

was observed in cultures expressing the Swedish mutation in the β -APP₇₅₁ form of the precursor (Fig. 2). Similar increases in A β were found by phosphor-imager analysis of the 4K A β band in gels of R1280 immunoprecipitates. This method further demonstrated that p3 was decreased several fold in the medium of the Swedish transfectants. We found that 293 and CHO cells stably transfected with the Swedish mutation in β -APP₇₅₁ also showed marked increases in A β in conditioned medium (data not shown).

To examine the mechanism for this increased A β production, we investigated the effects of each of the two mutations separately. Cells expressing β -APP₆₉₅ containing the 596 Met $\rightarrow$ Leu substitution had increased levels of A β in their medium, whereas cells expressing the 595 Lys $\rightarrow$ Asn substitution had levels similar to wild-type transfectants (Fig. 3, lanes 2-5). This finding suggests that the mutation at 596 is responsible for producing more proteolytic cleavage at the Leu-Asp peptide bond than at the normal Met-Asp bond. It is possible that the Lys $\rightarrow$ Asn switch at position 595 may further enhance the cleavage when coupled with the Met $\rightarrow$ Leu substitution. Next we created a DNA construct combining the Swedish double mutation with a deletion of almost all of the cytoplasmic domain of β -APP, thus removing the Asn-Pro-Xxx-Tyr consensus sequence¹⁵ for coated pit-mediated internalization of cell-surface proteins which may direct the reinternalization of β -APP from the cell surface and its targeting to late endosomes/lysosomes¹⁶. Cells transfected with this construct still produce substantially more A β in their media

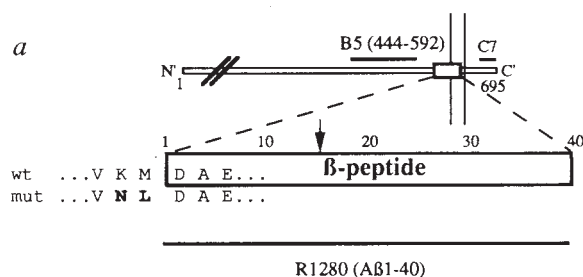


FIG. 1 Immunoprecipitation of A β from medium of cells transfected with wild-type or Swedish mutant β -APP₆₉₅ constructs. **a**, Schematic of β -APP and the A β region within it. Solid horizontal lines represent the regions against which antibodies B5, C7 and R1280 were raised; vertical lines designate the single membrane-spanning region of β -APP. Arrow indicates the constitutive secretory cleavage site of β -APP¹³. The amino-acid sequences of wild-type and mutant β -APP around the N terminus of A β are indicated. **b**, Conditioned medium of radiolabelled kidney 293 cells transiently transfected with either β -APP₆₉₅ (ref. 10) (lane 2) or β -APP₆₉₅ KM-NL (595 Lys $\rightarrow$ Asn + 596 Met $\rightarrow$ Leu) (lane 3) were immunoprecipitated with R1280. ¹²⁵I-labelled synthetic A β_{1-40} (ref. 6) was run as a size marker on the same gel (lane 1). The A β and p3 bands⁶ are indicated by arrows. **METHODS.** A cDNA construct encoding β -APP₆₉₅ KM-NL as described⁵ was designed by replacing the 26-base-pair *Bgl*III-*Eco*RI fragment of β -APP₆₉₅ with the annealed oligonucleotides GATCTCTGAAGTGAATCTGGATGCAG and AATTCTGCATCCAGATTCACCTCAGA. Transient transfections into 293 cells were performed using lipofectin (Gibco, BRL) as described by the manufacturer. Metabolic labelling with ³⁵S-methionine and immunoprecipitations with R1280 were done as described⁵. The immunoprecipitates were separated on a 10-20% Tris-Tricine gel, and autoradiographed for 9 days. The high-molecular-weight band (~218K) in lanes 2 and 3 is nonspecific and is variably precipitated with preimmune serum⁶.

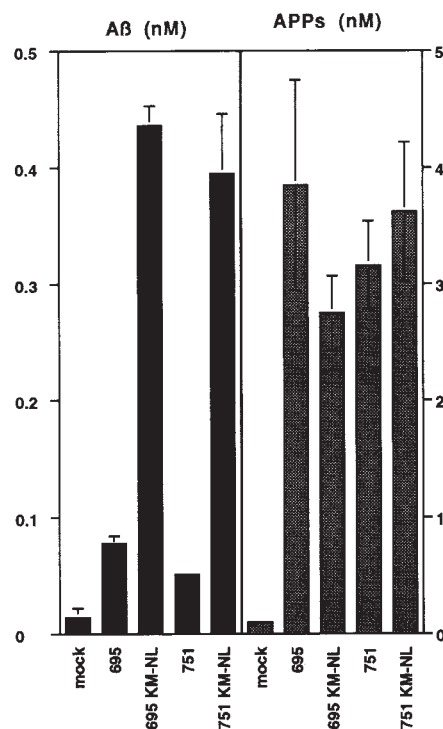


FIG. 2 Quantitation of A β (left panel) and APP_s (right panel) in conditioned medium of transiently transfected 293 cells using two distinct sandwich ELISAs. Each column represents the mean of four transfections, with the exception of the mock column which is based on three cultures. Error bars indicate the standard deviation. For columns without error bars, the standard deviation was less than 0.01.

METHODS. The sandwich ELISA for A β was done as before⁷ with monoclonal antibodies 266 (against A β_{1-28}) and 6C6 (against A β_{1-16}). The sandwich ELISA for APP_s was similarly constructed with affinity-purified polyclonal antibodies, using capture antibody B3 against a bacterial fusion protein of β -APP₂₀₋₃₀₄ and reporter antibody B5 (biotinylated) against a bacterial fusion protein of β -APP₄₄₄₋₅₉₂ (ref. 12) (β -APP₆₉₅ numbering). For each ELISA, increasing amounts of purified synthetic A β_{1-40} or purified APP_s from the conditioned medium of 293 cells transfected with β -APP₆₉₅ were used to construct a standard curve.

than wild-type transfectants, but levels of p3 are increased (Fig. 3, lanes 6 and 7). This result indicates that the effect of the Swedish mutations does not require an intact cytoplasmic domain and that generation of A β is unlikely to require processing of β -APP in late endosomes/lysosomes, a conclusion consistent with other recent results¹⁷.

The findings reported here provide, to our knowledge, the first experimental evidence that point mutations in the β -APP gene found in FAD kindreds can result directly in increased generation of the A β peptide. To date, six β -APP missense mutations have been identified in families with autosomal dominant Alzheimer's disease¹⁻⁵. In those cases in which the neuropathological phenotype has been documented, there is an early onset of progressive and severe β -amyloid deposition, particularly in the cerebral cortex. In the case of the Swedish family with the double mutations, the onset of clinical symptoms occurs at around 55 years of age⁵. Two different assays used here consistently demonstrate a marked increase in A β production by cells expressing the mutant precursor. A comparable decrease occurs in the production of the p3 peptide that begins at A β ₁₇₋₁₈, the site of constitutive secretory cleavage of β -APP¹³. Several lines of evidence, including pulse-chase experiments, isolation of lysosomes, the use of drugs affecting protein processing, and the deletion of the putative internalization signal in the cytoplasmic domain, all suggest that p3 derives from the 10K C-terminal stub of β -APP after secretion of APP_s¹⁷. The increase in p3 seen in the Δ C constructs (Fig. 3, lanes 6 and 7) presumably derives from a decrease in reinternalization of these truncated β -APP molecules to lysosomes, with consequent increased secretory cleavage occurring at or near the cell surface.

The increase in A β and the corresponding decrease in p3 in the Swedish transfectants suggests that A β may arise after an alternative secretory cleavage that generates the A β N terminus. A shorter form of APP_s, which this alternative cleavage would produce, has recently been described¹⁸. The lack of attenuation of A β production by cells expressing the Swedish β -APP construct with a cytoplasmic deletion supports such an alternative secretory mechanism, rather than the generation of A β as a result of internalization of β -APP to an endosomal/lysosomal pathway, although we cannot yet rule out A β generation in

early endosomes. This issue is now being addressed by developing antibodies that specifically recognize the last few residues of the shorter APP_s form including the substituted amino acids at 595 and 596 (ref. 18), and establishing whether such a secreted form increases in parallel with the rise of A β in the medium of the Swedish transfectants.

The double mutation of β -APP₅₉₅ and β -APP₅₉₆ consistently raises the level of A β *in vitro*, and the substitution of Leu for Met at 596 appears to be mainly responsible for this effect. It is important to determine whether the mutations at residue 717 of β -APP₇₇₀ immediately following the A β region¹⁻³ and those within the A β region^{4,19} also raise A β production. It is likely that these and other mutations will be found to affect A β production to different degrees and by different mechanisms. Additional artificial mutations can also be examined *in vitro* to define the amino-acid specificities of the proteases that cleave at the N and C termini of A β . Analysis of the Swedish mutations demonstrates the advantage of measuring A β production *in vitro*⁶⁻⁸ in transfected or primary cells bearing a particular mutation as a route to elucidating the specific mechanisms of accelerated β -amyloidosis in some familial forms of Alzheimer's disease. It will be interesting to see whether this marked rise in A β production *in vitro* is reflected in a change in the levels of the peptide in the body fluids of the Swedish patients. Moreover, β -APP-transfected cells or endogenous cells (such as fibroblasts) taken from a patient with the mutation could prove to be useful as a screen to identify compounds that lower cellular A β production. □

Received 12 October; accepted 16 November 1992.

- Goate, A. *et al.* *Nature* **349**, 704-706 (1991).
- Chartier-Harlin, M. D. *et al.* *Nature* **353**, 844-846 (1991).
- Murrell, J., Farlow, M., Ghetti, B. & Benson, M. D. *Science* **254**, 97-99 (1991).
- Hendriks, L. *et al.* *Nature Genet.* **1**, 218-221 (1992).
- Mullan, M. *et al.* *Nature Genet.* **1**, 345-347 (1992).
- Haass, C. *et al.* *Nature* **359**, 322-325 (1992).
- Seubert, P. *et al.* *Nature* **359**, 325-327 (1992).
- Shoji, M. *et al.* *Science* **258**, 126-129 (1992).
- Kang, J. *et al.* *Nature* **325**, 733-736 (1987).
- Selkoe, D. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 7341-7345 (1988).
- Weidemann, A. *et al.* *Cell* **57**, 115-126 (1989).
- Oltersdorf, T. *et al.* *J. biol. Chem.* **265**, 4492-4497 (1990).
- Esch, F. S. *et al.* *Science* **248**, 1122-1124 (1990).
- Sisodia, S. S., Koo, E. H., Beyreuther, K., Unterbeck, A. & Price, D. L. *Science* **248**, 492-495 (1990).
- Chen, W. J., Goldstein, J. S. & Brown, M. S. *J. biol. Chem.* **265**, 3116-3123 (1990).
- Haass, C., Koo, E. H., Mellon, A., Hung, A. Y. & Selkoe, D. J. *Nature* **357**, 500-503 (1992).
- Haass, C., Hung, A. Y., Schlossmacher, M. G. & Selkoe, D. J. *J. biol. Chem.* (in the press).
- Seubert, P. *et al.* *Nature* (in the press).
- Levy, E. *et al.* *Science* **248**, 1124-1126 (1990).

ACKNOWLEDGEMENTS. We thank L. Bickerstaff, D. Lee and M. Lee for technical assistance. This work was supported by grants from the NIH, Athena Neurosciences, and the Foundation for Neurologic Diseases to D.J.S., the Max-Planck-Gesellschaft (Otto-Hahn-Preis) to M.C., and Merck Sharp and Dohme Research Laboratories to A.Y.H.

Projection domains of MAP2 and tau determine spacings between microtubules in dendrites and axons

J. Chen*, Y. Kanai*, N. J. Cowan† & N. Hirokawa*‡

* Department of Anatomy and Cell Biology, School of Medicine, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

† Department of Biochemistry, New York University, Medical Center, New York, New York 10016, USA

NEURONS develop a highly polarized morphology consisting of dendrites and a long axon. Both axons and dendrites contain microtubules and microtubule-associated proteins (MAPs) with characteristic structures¹. Among MAPs, MAP2 is specifically

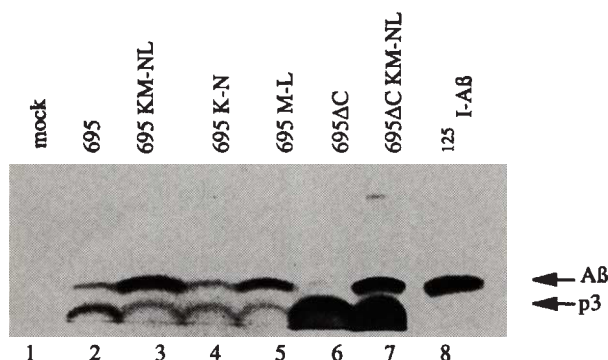


FIG. 3 Immunoprecipitation of A β from medium of cells transfected with various normal or mutant β -APP₆₉₅ constructs. Conditioned medium of radiolabelled kidney 293 cells transiently transfected with no DNA (lane 1), β -APP₆₉₅ (lane 2), β -APP₆₉₅ KM-NL (lane 3), β -APP₆₉₅ K-N (595 Lys $\rightarrow$ Asn; lane 4), β -APP₆₉₅ M-L (596 Met $\rightarrow$ Leu; lane 5), β -APP₆₉₅ Δ C (cytoplasmic deletion; lane 6) and β -APP₆₉₅ Δ C KM-NL (lane 7) were immunoprecipitated with R1280. ¹²⁵I-labelled synthetic A β (1-40) was run as a size marker on the same gel (lane 8). The A β and p3 bands⁶ are indicated by arrows.

METHODS. cDNA constructs encoding β -APP₆₉₅ K-N and β -APP₆₉₅ M-L were constructed as described for β -APP₆₉₅ KM-NL (see Fig. 1) using the annealed oligonucleotides GATCTCTGAAGTGAATATGGATGCAG, AATTC-TGCATCCATATTCACCTCAGA and GATCTCTGAAGTGAAGCTGGATGCAG. AATTC-TGCATCCAGCTTCACTTCAGA, respectively. A construct encoding β -APP₆₉₅ Δ C KM-NL was engineered from β -APP₆₉₅ Δ C (ref. 17) using the strategy described in Fig. 1. Transfections and R1280 immunoprecipitations were as for Fig. 1.

‡ To whom correspondence should be addressed.

expressed in dendrites² whereas MAP2C and tau are abundant in the axon²⁻⁵. But the influence of MAP2, MAP2C and tau on the organization of microtubule domains in dendrites versus axons is unknown. Both MAP2 and tau induce microtubule bundle formation in fibroblasts after transfection of complementary DNAs^{6,7}, and a long process resembling an axon is extended in Sf9 cells infected with recombinant baculovirus expressing tau^{8,9}. We have now expressed MAP2 and MAP2C in Sf9 cells in order to compare their morphology and the arrangement of their microtubules to that found in Sf9 cells expressing tau. We report here that the spacing between microtubules depends on the MAP expressed: in cells expressing MAP2, the distance is similar to that found in dendrites, whereas the spacing between microtubules in cells expressing MAP2C or tau is similar to that found in axons.

Insect ovarian Sf9 cells were infected with viruses containing recombinant sequences encoding the three major MAPs in mammalian neurons; MAP2 (1,828 amino acids)⁷, MAP2C (467

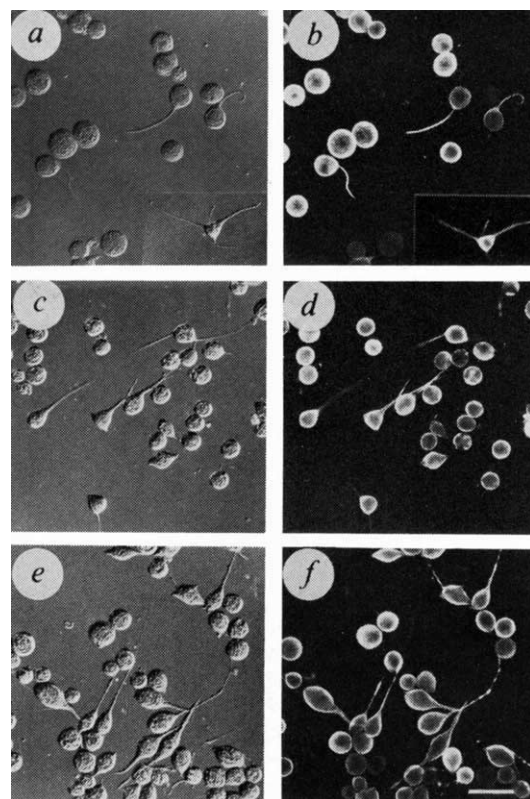


FIG. 1 Sf9 cells infected with recombinant baculovirus expressing MAP2 (a, b), MAP2C (c, d), or tau (e, f) examined using Nomarski optics (a, c, e) or immunofluorescence microscopy after staining with anti-MAP2 antibody (b, d) or anti-tau antibody (f). Scale bar, 50 μ m.

METHODS. Construction of recombinant baculovirus, cell culture, and baculovirus infection were done according to standard methods²⁴. 48h after infection, cells were stained with standard indirect immunofluorescence methods using either a monoclonal anti-MAP2 antibody²⁵ or an anti-tau antibody³ (Tau-1, a generous gift of L. I. Binder) after fixation for 15 min at 37 °C with 2% paraformaldehyde and 0.1% glutaraldehyde in PEM buffer (80 mM Pipes, 1 mM MgCl₂, 1 mM EGTA).

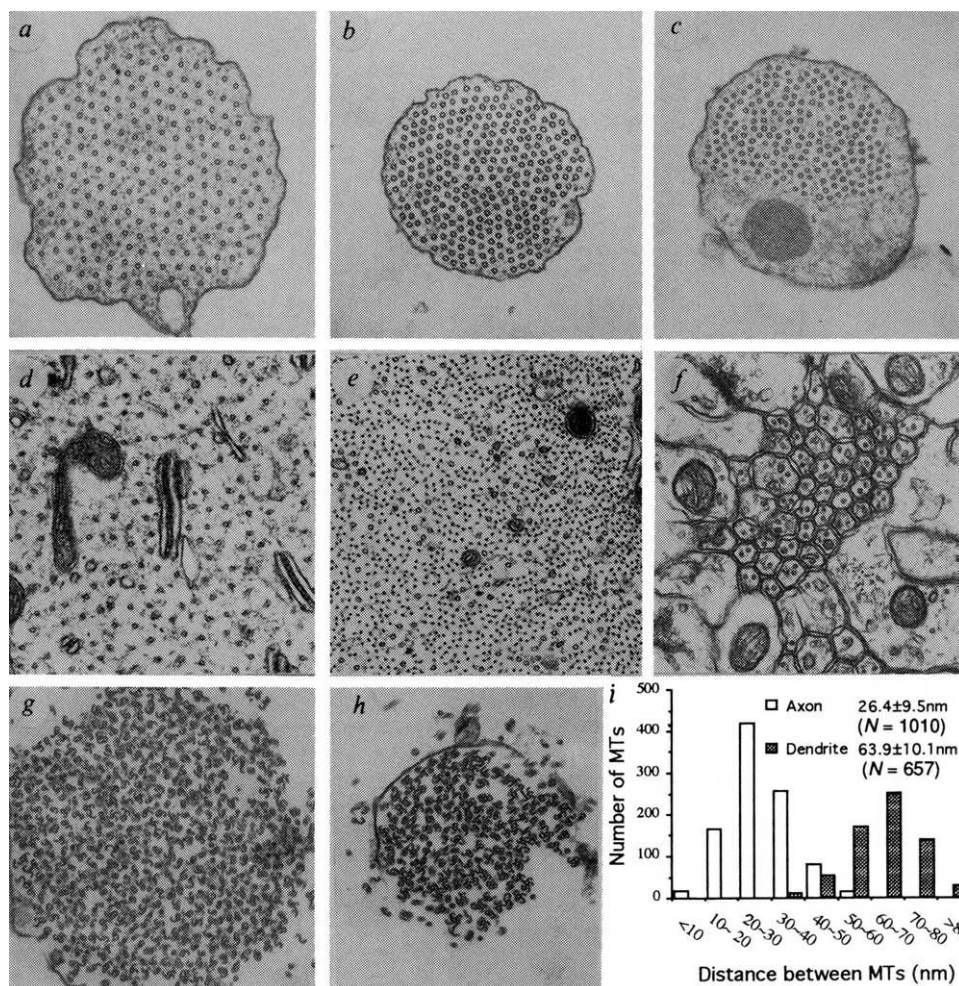


FIG. 2 Cross-section electron micrographs of MAP-induced processes in Sf9 cells, fixed as described (see methods) (a, MAP2; b, MAP2C; c, tau), or after 'hook formation' to determine the polarity of microtubules (g, MAP2; h, MAP2C). d, e, f, Cross-section electron micrographs of a Purkinje cell dendrite (d), a large myelinated axon in the spinal cord (e), and parallel fibre axons in rat cerebellum (f). i, Compilation of wall-to-wall distances between nearest adjacent microtubules in Purkinje cell dendrites and spinal cord axons. In axons, microtubules tend to form small fascicles surrounded by neurofilaments (as exemplified by encircled areas). Measurements were therefore made of the distance between nearest microtubules in those cases where there were no intervening neurofilaments or membranous organelles. Scale bar, 250 nm.

METHODS. Rat tissues or infected Sf9 cells were fixed with 2.5% glutaraldehyde and 2% paraformaldehyde in 0.1 M cacodylate followed by post-fixation with 1% OsO₄. Hook formations were done according to the procedure by Heidemann *et al.*²⁰.

amino acids)¹⁰ or tau (the latter as the largest (432 amino acids) isoform)⁶. After infection, the host cells synthesized these MAPs and one or multiple long processes were extended from the normally rounded cells (Fig. 1). Electron microscopy revealed microtubule bundles in all the processes. In serial thin sections, we found that the distances between adjacent microtubules were quite distinct among cells expressing MAP2, tau and MAP2C. The average wall to wall distance between adjacent nearest microtubules in microtubule bundles in the middle portion of processes induced by MAP2 was 61.7 ± 9.0 nm ($n = 158$) from four samples (Fig. 2a), whereas in MAP2C expressing cells the corresponding distance was 20.8 ± 3.6 nm ($n = 122$) from four samples (Fig. 2b) or 19.8 ± 4.1 nm ($n = 137$) from five samples in the case of tau-expressing cells (Fig. 2c).

We compared the microtubule organization in Sf9 cells expressing MAP2, MAP2C or tau with those in the dendrites and axons of normal neuronal cells. For this purpose we chose rat cerebellar Purkinje cell dendrites expressing abundant quantities of MAP2, parallel fibre axons and axons in the white matter of spinal cords expressing abundant quantities of tau. The distances between adjacent nearest microtubules in Purkinje cell dendrites, parallel fibre axons and white matter spinal cord axons were 63.9 ± 10.1 nm ($n = 657$) (Fig. 2d), 22.2 ± 9.8 nm ($n = 54$) (Fig. 2f) and 26.4 ± 9.5 nm ($n = 1,010$) (Fig. 2e), respectively. In axons of large diameter which contain abundant neurofilaments, microtubules tend to form small fascicles. We therefore compared the distribution of distances between nearest adjacent microtubules in Purkinje cell dendrites and spinal cord large axons (Fig. 2i) and found an average inter-microtubule spacing of 60–70 nm in dendrites and 20–30 nm in axons (Fig. 2i). Overall, the inter-microtubule spacing in dendrites corresponds very well with the distance between microtubules in MAP2-expressing Sf9 cells, whereas in axons the inter-micro-

tubule spacing is close to the distance between microtubules in tau or MAP2C-expressing Sf9 cells. Furthermore, quick-freeze deep-etch observation of microtubule bundles in infected Sf9 cells revealed that microtubules are crosslinked with each other through long (50–120 nm) and sometimes branching filamentous structures in MAP2-expressing cells (Fig. 3a, arrows) whereas the filamentous crossbridges are short and straight (15–30 nm) in MAP2C- or tau-expressing cells (Fig. 3b, c, arrows). Because the filamentous structures between microtubules look exactly like those between microtubules polymerized *in vitro* in the presence of MAP2 (Fig. 3d, arrows), MAP2C (ref. 11) or tau (Fig. 3e, arrows), respectively, it is reasonable to suppose that these filamentous structures are composed of MAP2, MAP2C or tau.

MAP2, MAP2C and tau all share a homologous C-terminal domain containing three (or more) repeated sequences that form part of the microtubule-binding domain^{6,7,12–18}. The principal difference between these molecules is in the sequence and extent of their N-terminal domains (projection domain)¹⁹: in the case of MAP2, this domain is 1,669 amino-acids long, whereas the corresponding domains in MAP2C and tau are of 308 and 242 amino acids, respectively. It seems likely, therefore that the different length of crossbridges between microtubules observed *in vivo* and *in vitro* in this study is a result of differences in the N-terminal rod domains.

The polarity of microtubules in axons is uniform with the plus end pointing towards the cell periphery²⁰, whereas in dendrites the polarity is mixed²¹. We therefore asked whether the major MAP expressed in each compartment has a role in determining polarity. We found that in MAP2- and MAP2C-induced microtubule bundles, most microtubules are of uniform polarity with the plus end pointing distally (95.3% in MAP2-expressing cells (Fig. 2g) and 94.5% in MAP2C-expressing cells (Fig. 2h)).

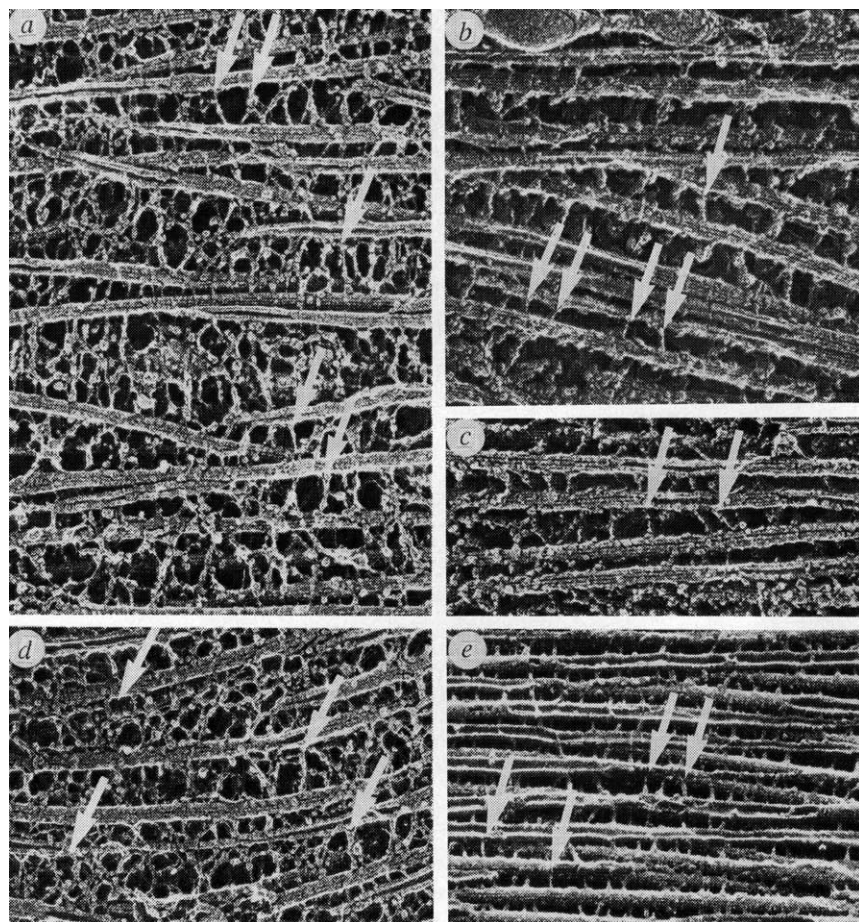


FIG. 3 Quick-freeze, deep-etch electron microscopy of processes in infected Sf9 cells (a, MAP2; b, MAP2C; c, tau), and microtubules copolymerized *in vitro* with either MAP2 (d) or tau (e). Scale bar, 100 nm.

METHODS. Infected cells were extracted for 10 min at room temperature in 0.05% saponin, 5 μ M taxol 2 mM GTP in PEM buffer, and quick-frozen as described previously²⁶. *In vitro* copolymerization of microtubules with MAP2 or tau was done as described elsewhere^{27,28}.

Microtubules in tau-expressing cells also have the same polarity⁹. Therefore, the polarity of microtubules in dendrites must be determined by factors other than interaction with MAP2 alone.

The mechanism underlying the bundling of microtubules by MAP2 and tau is likely to be complex^{22,23}. Nonetheless, the present study clearly shows that in cells expressing a single brain MAP species, crossbridges between microtubules are formed whose length varies as a function of the length of the N-terminal arm. Furthermore, the inter-microtubule spacing also depends on the length of the N-terminal arm. Thus, the projection domains of these molecules seem likely to be key determinants of the spacing between microtubules in dendrites and axons. □

Received 6 August; accepted 23 October 1992.

- Hirokawa, N. in *Neuronal Cytoskeleton* (eds R. D. Burgoyne) 5–74 (Wiley Liss, New York, 1991).
- Bernhardt, R. & Matus, A. *J. comp. Neurol.* **226**, 203–231 (1984).
- Binder, L. I., Frankfurter, A. & Rebhun, L. I. *J. Cell Biol.* **101**, 1371–1378 (1985).
- Matus, A. *Rev. Neurosci.* **11**, 29–44 (1988).
- Nunez, J. *Trends Neurosci.* **11**, 477–479 (1988).
- Kanai, Y. *et al. J. Cell Biol.* **109**, 1173–1184 (1989).
- Lewis, S. A., Ivanov, I. E., Lee, G. H. & Cowan, N. *J. Nature* **342**, 498–505 (1989).
- Knops, J. *et al. J. Cell Biol.* **114**, 725–733 (1991).
- Baas, P. W., Pienkowski, T. P. & Kosik, K. S. *J. Cell Biol.* **115**, 1333–1344 (1991).
- Uemeyama, T., Okabe, S., Kanai, Y. & Hirokawa, N. *J. Cell Biol.* (in the press).
- Takeuchi, M., Hisanaga, S., Uemeyama, T. & Hirokawa, N. *J. Neurochem.* **58**, 1510–1516 (1992).
- Goedert, M., Wischik, C. M., Crowther, R. A., Walker, J. E., & Klug, A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4051–4055 (1988).
- Lee, G., Cowan, N. & Kirschner, M. W. *Science* **239**, 285–288 (1988).
- Aizawa, H. *et al. J. Biol. Chem.* **263**, 7703–7707 (1988).
- Goedert, M., Spillantini, M. G., Potier, M. C., Ulrich, J. & Crowther, R. A. *EMBO J.* **8**, 393–399 (1989).
- Kosik, K. S., Orecchio, L. D., Bakalis, S. & Neve, R. L. *Neuron* **2**, 1389–1397 (1989).
- Lee, G., Neve, R. L. & Kosik, K. S. *Neuron* **2**, 1615–1624 (1989).
- Papadimitrakopoulou, A., Doll, T., Tucker, R. P., Garner, C. C. & Matus, A. *Nature* **340**, 650–652 (1989).
- Vallee, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3206–3210 (1980).
- Heidemann, S. R., Landers, J. M. & Hamborg, M. A. *J. Cell Biol.* **91**, 661–665 (1981).
- Baas, P. W., Deitch, J. S., Black, M. M. & Banker, G. A. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8335–8339 (1988).
- Kanai, Y., Chen, J. & Hirokawa, N. *EMBO J.* **11**, 3953–3962 (1992).
- Wille, H. *et al. J. struct. Biol.* **108**, 49–61 (1992).
- Webb, N. R. & Summers, M. D. *MAXBAC™: Baculovirus Expression System. A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures* (Invitrogen, San Diego, 1990).
- Sato-Yoshitake, R., Shiomura, Y., Miyasaka, H. & Hirokawa, N. *Neuron* **3**, 229–238 (1989).
- Hirokawa, N. *J. Cell Biol.* **94**, 129–142 (1982).
- Hirokawa, N., Shiomura, Y. & Okabe, S. *J. Cell Biol.* **107**, 1449–1461 (1988).
- Hirokawa, N., Hisanaga, S. & Shiomura, Y. *J. Neurosci.* **8**, 285–306 (1988).

ACKNOWLEDGEMENTS. This work was supported by a grant-in aid for special project research by the Ministry of Education Science and Culture of Japan to N.H. and a HFSP grant to N.J.C. and N.H.

The spatial arrangement of cones in the primate fovea

J. D. Mollon* & J. K. Bowmaker†

* Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, UK

† Institute of Ophthalmology, University of London, Bath Street, London EC1V 9EL, UK

THE retinae of Old World primates contain three classes of light-sensitive cone, which exhibit peak absorption in different spectral regions^{1–4}. But how are the different types of cone arranged in the hexagonal mosaic of the fovea? This question has often been answered with artists' impressions^{5–7}, but never with direct measurements. Staining for antibodies specific to the short-wave photopigment has revealed a sparse, semiregular array of cones⁸; but nothing is known about the arrangement of the more numerous long- and middle-wave cones. Are they randomly distributed, with chance aggregations of one type, as Hartridge postulated in these columns nearly 50 years ago^{9,10}? Or do they exhibit a regular alternation, recalling the systematic mosaics seen in some non-mammalian species^{6,11}? Or, conversely, is there positive clumping of particular cone types, as might be expected if local patches of cones were descended from a single precursor cell? We have made direct microspectrophotometric measurements of patches of foveal retina from Old World monkeys, and report here that the distribu-

tion of long- and middle-wave cones is locally random. These two cone types are present in almost equal numbers, and not in the ratio of 2:1 that has been postulated for the human fovea.

Post-mortem tissue was made available to us from talapoin monkeys (*Cercopithecus talapoin*) that had been used in unrelated endocrinological studies. When the fresh retina was dissected under dim red light, the fovea was usually visible as a dark spot. A retinal fragment that included the fovea was mounted on a slide, was divided by two or three strokes of a razor blade, and was then partially dispersed by gentle pressure on a cover slip. On occasion we were able to produce small fragments of intact retina that were suitable for microspectrophotometry and yet were extensive enough to answer the present question. The absorbance spectra of individual receptors were measured with a modified dual-beam Liebman microspectrophotometer under computer control^{12–14}; the numerical apertures of the condenser and objective were 0.4 and 0.9, respectively⁴. With the help of an infrared converter, the measuring beam (2 μ m square cross-section) was aligned so as to pass axially through a given cell while the reference beam passed either outside the retinal fragment or in a space between photoreceptors. The spectrum was scanned from 750 to 370 nm and back to 750 nm. To minimize the effects of bleaching, only one absorption spectrum was usually obtained from a given cell, but two independent estimates of the baseline absorbance spectrum were always obtained by arranging both beams to pass outside the cell. In the case of cells that absorbed at short wavelengths, we checked that the pigment was photolabile by repeating the absorbance measurements after bleaching the cell with white light for 5 min.

We were able to measure 30–40 receptors in a given patch. Although microspectrophotometric measurements are usually made by passing the beam transversely through the outer segment¹⁵, our measurements of cells in these foveal patches were necessarily axial; the absorption spectra may thus be subject to minor distortion by inert pigments in the inner segments⁴, by waveguide effects, or by scattering in other retinal tissue. Nevertheless, we were always able to assign a cone to one of three categories without ambiguity. In Fig. 1 we plot the wavelength of peak sensitivity ($\lambda_{\max}$) for each of the long- and middle-wave cones recorded in the present study, including those used to construct Fig. 2: there is a clear separation of cones into two distinct classes and so we can be confident that our axial records are adequate for the present purpose of classification.

In Fig. 2 we show at the upper right a typical fragment of foveal retina from one of our preparations. The solid outline encloses the subset of the array that we were able to characterize

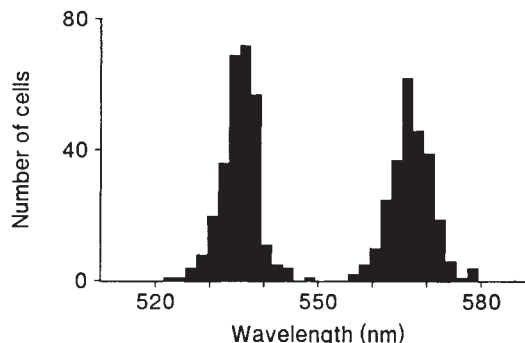
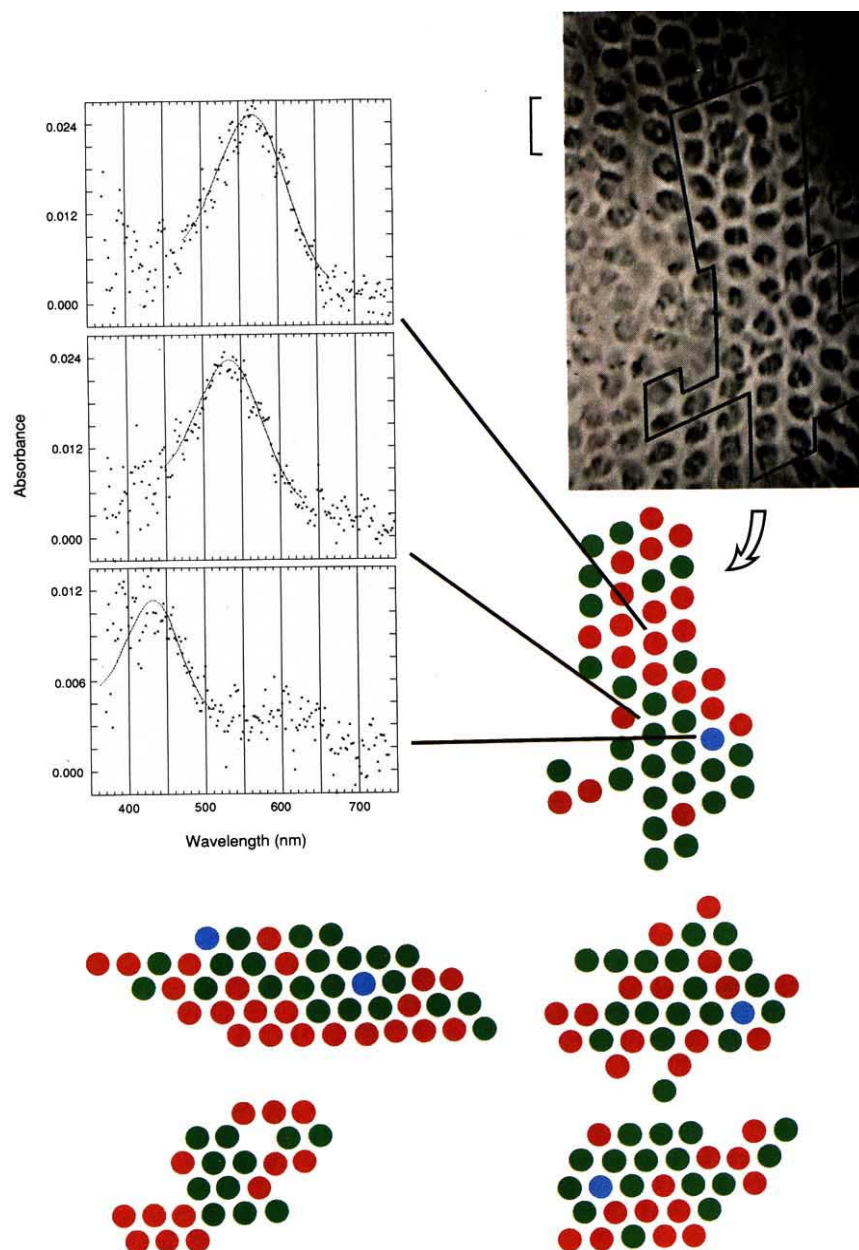


FIG. 1 The distribution of $\lambda_{\max}$ values for all middle- and long-wave cones in our sample, including those used in the construction of the schematic matrices of Fig. 2. For each individual record, the $\lambda_{\max}$ was estimated by fitting a template to the long-wave limb of the absorbance spectrum^{4,14}. Note that we have included in this sample some records that would not pass our standard criteria⁴ if our purpose were to establish the exact form of the mean absorbance spectrum. Nevertheless, the records fall into two distinct groups and there is no ambiguity in classifying individual cells as middle- or long-wave.

FIG. 2 The photograph at the upper right is of a fragment of fresh retina focused at the level of the outer segments and recorded on video by means of an infrared television camera permanently mounted on the microspectrophotometer. Scale bar, 10 μm . The solid outline encloses the area within which we measured each individual receptor. Immediately below the photograph, we show schematically the corresponding array of cones: long-, middle- and short-wave cones are represented by red, green and blue discs, respectively. For the three cones indicated, we show to the left their individual absorbance spectra, which are typical of those on which we base our classification; the solid curves fitted to the data are templates based on the Dartnall nomogram^{1,3} displaced on a log frequency abscissa⁴. The lower part of the figure shows four other retinal patches that have been similarly analysed.



fully. To the left are shown examples of individual absorbance spectra for the three types of cones: the λ_{max} values lie close to 565, 535 and 430 nm, as is typical for Old World primates¹⁻⁴. On the basis of similar spectra, we assigned each of a sample of 183 cones to one of three classes and so derived the matrices shown symbolically in the remainder of Fig. 2. The patches shown are drawn from five eyes from four female monkeys.

It is at once clear from Fig. 2 that the long- and middle-wave cones of the primate retina do not form the regular, systematically alternating array that has sometimes been postulated. Clumping of cones of the same type is apparent, but is this more than the chance aggregation that arises in any random distribution^{9,16}? For the five patches shown, we have counted the transitions between two middle-wave cones, between two long-wave cones, and between cones of different types. The counts were made in one, arbitrary, direction (corresponding to the long axis of the fragment) and transitions involving short-wave cones were excluded. A χ^2 test showed that the frequencies of the three possible types of transition did not differ significantly from the values expected if the identity of a cone were independent of the identity of its neighbours and reflected only the absolute

probability of that cone type ($\chi^2 = 1.22$, d.f. = 2). If there is any bias, it is towards aggregation rather than alternation.

Our finding that the foveal matrix is random rather than ordered has implications for the nature of the post-receptoral channels that subserve spatial and chromatic resolution. It would not be possible for the same midget ganglion cells to be optimized for both functions. Chromatic specificity of centre and surround inputs¹⁷ would require that the ganglion cell's opposing inputs are not always drawn from cones that are nearest neighbours. Conversely, if spatial resolution is optimized in the sense that opposing inputs are drawn from neighbouring cones, then the surround inputs cannot always be drawn from a single class of cone¹⁸. In either case, the random clustering of cones of one type must constrain colour discrimination at high spatial frequencies, and also has implications for chromatic aliasing⁷.

Of the 183 cones in our sample of intact retinal patches, long-wave, middle-wave and short-wave cones are present in the ratios 86:92:5 (47%, 50%, 3%). These records from intact patches are a subset of a total of 557 cones recorded from four male and six female talapoin monkeys. If we include all records (from intact patches and from isolated receptors) the corres-

ponding ratios are 256:289:12 (46%, 52%, 2%). The rarity of short-wave cones is concordant with earlier estimates, using direct and indirect means of identifying these receptors^{1-4,8,19,20}, but we have not ourselves seen any indication that the short-wave cones occur as interruptions to the regular foveal matrix, as has sometimes been reported²¹. The ratio of long- to middle-wave cones ranges from 1.6 to 0.42 for samples from individual animals and for the total sample has a value (0.89) that does not differ significantly from unity ($\chi^2 = 1.87$; d.f. = 1). This value is consistent with our earlier results for a number of catarrhine species⁴ and with the results of Hárosi³ and of Baylor, Nunn and Schnapf² for macaques, but it is extremely unlikely ($\chi^2 = 95$, d.f. = 1, $P < 0.001$) that our sample could be drawn from a

population with a ratio of 2:1, the ratio that has been proposed for man on the basis of psychophysical measurements^{22,23}.

If this species difference proves to be real, two questions arise. First, what ecological factors have led to a more balanced proportion of long- and middle-wave cones in monkeys than in man? (Is it that good colour vision at small visual angles has been more important for frugivorous monkeys than for our own recent ancestors²⁴?) Second, what is the molecular genetic mechanism that in monkeys leads to a near 1:1 ratio but in man to a strong bias in favour of long-wave cones? If we knew the answer to this question, we should be close to understanding the clinically important issue of how the expression of opsin genes is controlled. □

Received 10 August; accepted 5 October 1992.

1. Bowmaker, J. K., Dartnall, H. J. A. & Mollon, J. D. *J. Physiol.* **298**, 131-143 (1980).
2. Baylor, D. A., Nunn, B. J. & Schnapf, J. L. *J. Physiol.* **390**, 145-160 (1987).
3. Hárosi, F. *J. gen. Physiol.* **89**, 717-743 (1987).
4. Bowmaker, J. K. *et al. J. exp. Biol.* **156**, 1-19 (1991).
5. Walraven, P. L. *Vision Res.* **14**, 1339-1343 (1974).
6. Lythgoe, J. N. *The Ecology of Vision* (Oxford Univ. Press, Oxford, 1979).
7. Williams, D. R. *et al. in From Pigments to Perception* (eds Valberg, A. & Lee, B. B.) (Plenum, New York, 1991).
8. Curcio, A. C. *et al. J. comp. Neurol.* **312**, 610-624 (1991).
9. Hartridge, H. *Nature* **153**, 45-46 (1944).
10. Hartridge, H. *Nature* **157**, 482 (1946).
11. Bowmaker, J. K. & Kunz, Y. W. *Vision Res.* **27**, 2101-2108 (1987).
12. Liebman, P. A. & Entine, G. *J. opt. Soc. Am.* **54**, 1451-1459 (1964).
13. Knowles, A. & Dartnall, H. J. A. *The Photobiology of Vision* (Academic, London, 1977).

14. Mollon, J. D., Bowmaker, J. K. & Jacobs, G. H. *Proc. R. Soc. B* **222**, 373-399 (1984).
15. Dobelle, W. H., Marks, W. B. & MacNichol, E. F. Jr. *Science* **168**, 1508-1510 (1969).
16. Hartridge, H. *Phil. Trans. R. Soc. B* **232**, 519-671 (1947).
17. Reid, R. C., Shapley, R. M. *Nature* **356**, 716-718 (1992).
18. Lennie, P. & Haake, P. W. in *Computational Models of Visual Processing* (eds Landy, M. S. & Movshon, J. A.) 71-82 (MIT Press, Cambridge, MA, 1991).
19. Marc, R. E. & Sperling, H. G. *Science* **196**, 454-456 (1977).
20. de Monasterio, F. M., Schein, S. J. & McCrane, E. P. *Science* **213**, 1278-1281 (1981).
21. Anhalt, P. K., Kolb, H. & Pflug, R. J. *comp. Neurol.* **255**, 18-34 (1987).
22. Cicerone, C. M. & Neger, J. L. *Vision Res.* **29**, 115-128 (1989).
23. Pokorny, J., Smith, V. C. & Wesner, M. F. in *From Pigments to Perception* (eds Valberg, A. & Lee, B. B.) (Plenum, New York, 1991).
24. Mollon, J. D. *J. exp. Biol.* **146**, 21-38 (1989).

ACKNOWLEDGEMENTS. We thank M. Webster and B. C. Regan for discussion. Supported by the MRC and by Fight for Sight.

A cell line that can induce thymocyte positive selection

Patrice Hugo*, John W. Kappler*, Dale I. Godfrey† & Philippa C. Marrack*

* Howard Hughes Medical Institute, National Jewish Center for Immunology and Respiratory Medicine, 1400 Jackson Street, Denver, Colorado 80206, USA

† DNAX Research Institute of Molecular and Cellular Biology, Palo Alto, California 94304, USA

THE thymus positively selects thymocytes that bear T-cell receptors which recognize antigen presented by self major histocompatibility complex (MHC) proteins^{1,2}. Positive selection is usually driven by MHC products on radiation-resistant cortical epithelial cells³⁻⁵. It is unknown whether positive selection is mediated by all thymic epithelial cells or by some specialized subsets⁴. Here we introduce an H-2^b-expressing thymic epithelial cell line into the thymuses of lethally irradiated H-2^k animals reconstituted with H-2^{b/k} F₁ BM or fetal liver cells. I-A^b-restricted T cells are found in these animals, demonstrating that selection occurs on the introduced epithelial cells.

Long-term thymic cultures were established from C57BL/6 (H-2^b) mice. One clone, 2E4, was chosen on the basis of its reactivity with anti-cytokeratin antibodies. 2E4 cells were constitutively I-A^b-negative and bore low amounts of H-2K^b. High levels of both molecules were detected on 2E4 after interferon- γ (IFN- γ) treatment, an effect previously observed with other cultured thymic epithelial cells⁶. 2E4 cells did not express markers found on macrophages and dendritic cells, for example CD45, CD11a/18, HSA, CD11b/18, MAC-2, F4/80, and Fc γ RII/III. As for most cultured thymic epithelial lines, attempts to determine whether the 2E4 clone was derived from the cortex or the medulla were inconclusive.

2E4 cells pretreated with IFN- γ presented the superantigen staphylococcal enterotoxin B, which does not require processing, to the DO-11.10 T-cell hybridoma (Fig. 1). 2E4 cells presented chicken ovalbumin efficiently to BO-97.11, but very poorly to DO-11.10. But DO-11.10 was stimulated by the ovalbumin

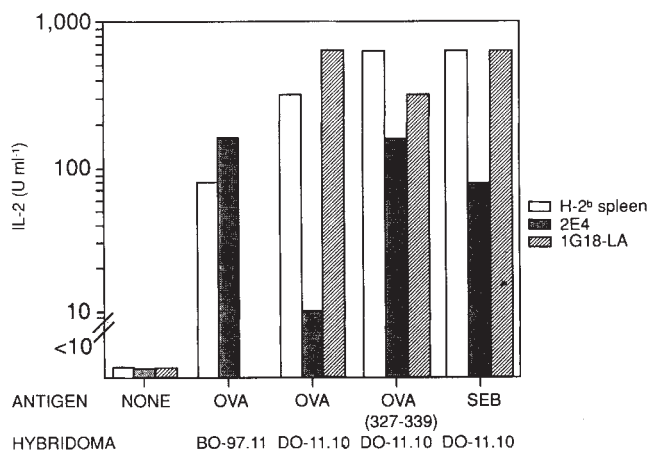


FIG. 1 Antigen presentation by 2E4 epithelial cells. The capacity of 2E4 cells (H-2^b) to present chicken ovalbumin (OVA), OVA peptide 327-339, and staphylococcal enterotoxin B (SEB) to T-cell hybridomas was compared to that of fresh splenocytes from C57BL/10 mice (H-2^b) or that of the thymic H-2^d macrophage cell line 1G18LA.

METHODS. One of us (D.I.G.) established long-term cultures from thymuses of 6-8-week-old C57BL/6 female mice in α -MEM supplemented with 10% fetal calf serum, 10% horse serum, 5×10^{-5} M 2-mercaptoethanol and 5×10^{-7} M hydrocortisone. Cultures were passaged for several months and the 2E4 clone was obtained by two rounds of cloning and thereafter cultured in complete medium without hydrocortisone. Antigen presentation was assayed essentially as described⁷. Briefly, 2E4 and 1G18LA cells (gift from A. Zlotnik⁷) were plated in microculture plates (2.5×10^4 per well) and incubated in the presence of IFN- γ (200 U ml^{-1}) for 48 h. Cells were washed and OVA-specific T-cell hybridomas BO-97.11 (restricted by H-2^b) or DO-11.10 (restricted by H-2^d; TCR V β 8.2⁺) were added at 10^5 per well. Fresh splenocytes (10^6 per well) from C57BL/10 mice were also used as a source of antigen-presenting cells. Cultures were incubated for 24 h with 1 mg ml^{-1} OVA (Sigma), $30 \text{ } \mu\text{g ml}^{-1}$ OVA peptide 327-339 (Macromolecular Resources) or $10 \text{ } \mu\text{g ml}^{-1}$ SEB (Sigma). Interleukin 2 (IL-2) was assayed using HT-2 (ref. 21) and expressed as relative units per ml. No IL-2 was detected without preincubation of the 2E4 or 1G18LA cells with IFN- γ .

peptide 327–339. Therefore the 2E4 cells processed and presented ovalbumin, but not in a form recognized by DO-11.10. This limited ability to process ovalbumin appears to be a feature of thymic epithelial cells⁷. Nonetheless, expression of functional I-A molecules was inducible on 2E4 cells and these cells could process and present antigens.

The 2E4 cells were injected intrathymically to assess their capacity to mediate positive selection. We first tested whether the 2E4 cells would colonize the thymic cortex, where positive selection occurs^{8–10}. 2E4 cells were marked before injection with PKH26, a red fluorescent membrane dye. Eight days later, recipient thymuses were collected and were stained with a fluorescent (green)-conjugated anti-keratin antibody. Because 2E4 cells express cytokeratin (Fig. 2a), they were visualized as yellow (green + red) after double exposure. They were found in the cortex (Fig. 2b), and the medulla (not shown) of recipient thymus, indicating that they penetrated the cortex and suggesting that they can persist for a period long enough to mediate positive selection.

To test this idea directly, B10.BR (H-2^k) or (C57BL/10 × B10.BR)F₁ (H-2^{b/k}) (F₁) mice were irradiated and reconstituted with F₁ BM cells. Three weeks later, when the rate of positive selection is maximal^{11,12}, 2E4 (H-2^b) cells were injected into thymic lobes of reconstituted mice. Control animals were injected with saline. Reconstitution with F₁ BM cells induced tolerance to H-2^b, thus preventing rejection of 2E4 cells, and allowed priming for measurement of H-2^b restricted T-cell activity. At the time of injection, the 2E4 cells were I-A⁺. Four weeks after introduction of 2E4 (H-2^b) cells into F₁ → B10.BR (H-2^k) chimaeras, thymus sections were prepared and stained for I-A^b

(green) and cytokeratin (red). Expression of I-A^b on 2E4 epithelial cells was visualized as yellow on a double exposure (Fig. 2c). This suggests I-A^b has been induced (although at low levels) *in vivo* on 2E4 cells. It is known that interactions between thymus epithelial cells and T-cell precursors induce class II on the former cells^{13,14}, perhaps as a result of IFN- γ production by the latter^{14–17}.

Four weeks after intrathymic injection, mice were immunized with (Tyr-Glu)-poly(D,L-Ala)-poly(Lys) ((T,G)-A--L), an antigen recognized by T cells in association with I-A^b (2E4 haplotype), but recognized poorly with I-A^k (host thymus haplotype) (ref. 18; Fig. 3a, upper panel). The former response was blocked at 90% with anti-I-A^b antibodies (P.H., unpublished observation). One week later, T cells from these animals were assessed for their response to (T,G)-A--L presented by H-2^b splenocytes. As shown in Fig. 3a (lower panel), and as expected, T cells from F₁ → F₁ chimaeras responded well, but T cells from F₁ → B10.BR chimaeras responded very poorly to (T,G)-A--L presented by H-2^b splenocytes. This weak response suggests that these animals may have had some I-A^b-restricted T cells and could reflect the ability of BM-derived cells to induce positive selection, albeit inefficiently^{19,20}. On the other hand, T cells from F₁ → B10.BR chimaeras inoculated intrathymically with 2E4 cells responded to (T,G)-A--L at an intermediate level (Fig. 3a, lower panel). Although this increased response was small, it was reproducible; results were similar in three independent experiments (Fig. 3b). Such a weak increase was expected because host (H-2^k) thymic epithelial cells should continue to select positively and would probably, because of numerical and topographical advantages, be much more efficient than the

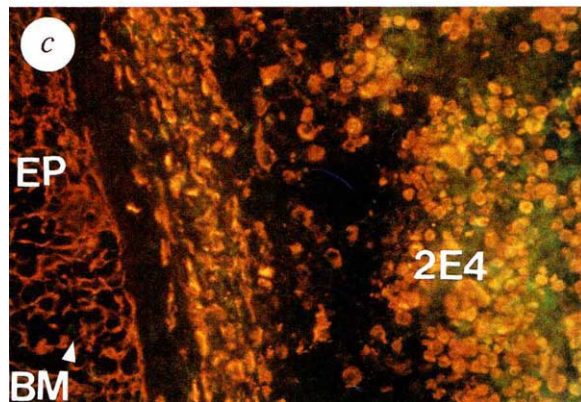
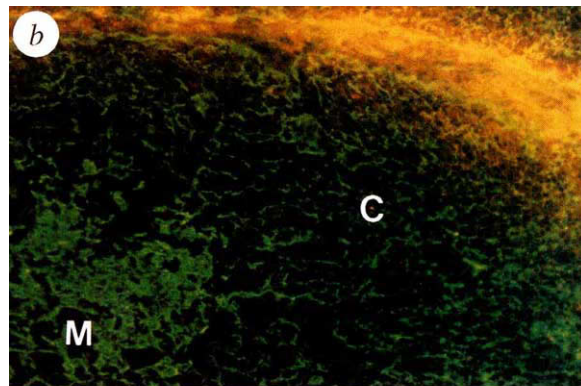
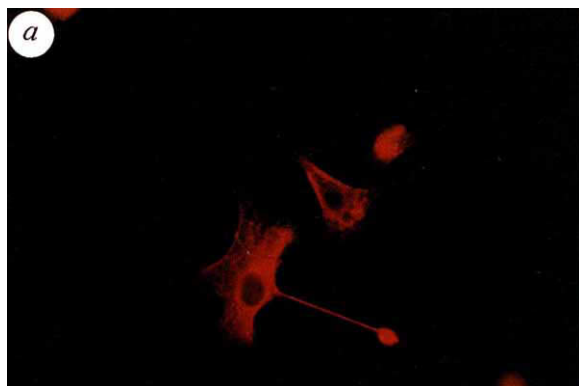


FIG. 2 Thymic localization and I-A induction after intrathymic injection of 2E4 cells. Cytokeratin expression was examined in 2E4 cells (a, magnification 1000 $\times$). 2E4 cells were marked with the fluorescent (red) membrane dye PKH26 and injected intrathymically. Their localization was determined 8 days later on frozen thymic sections with a keratin counterstain (green). The photograph was double-exposed to allow simultaneous visualization of red and green staining, hence the yellow colour indicates cells positive for both markers (b, magnification, 500 $\times$). *In vivo* induction of I-A^b on 2E4 cells was assessed 4 weeks after their injection into F₁ → B10.BR chimaeras by staining for anti-I-A^b (green) and keratin (red) (c, magnification 500 $\times$). The photograph shows endogenous epithelium (EP: red) as I-A^b keratin⁺, BM-derived cells (BM: green) as I-A^b keratin⁺, and a majority of injected 2E4 cells (2E4: yellow) as I-A^b keratin⁺. A minority of green-stained cells may be 2E4 cells that have lost the epitope recognized by the anti-keratin antibody, as seen with 2E4 cultures.

METHODS. 2E4 cells were cultured on Multitest Slides (Flow Laboratories), fixed with cold acetone for 2 min and stained with a rabbit polyclonal anti-keratin antibody (Dako) revealed with a rhodamine-conjugated goat anti-rabbit antibody (FisherBiotech) (a). 2E4 cells were incubated with the red membrane dye PKH26-GL (4 $\times$ 10⁻⁶ M) (Zynaxis Cell Science) for 5 min and washed in saline to remove free dye. Intrathymic injections were done as described²² with 10⁷ 2E4 cells in 20 μ l saline. Later, thymuses were collected and snap-frozen in liquid nitrogen. Thin sections (6–8 μ m) were

fixed with cold acetone and stained with the anti-keratin antibody and revealed with a goat anti-rabbit antibody (FisherBiotech) conjugated to fluorescein (b), or rhodamine when the fluorescein-conjugated anti-I-A^b antibody (clone D3.137.5) was used (c). Background fluorescence was determined using irrelevant antibodies.

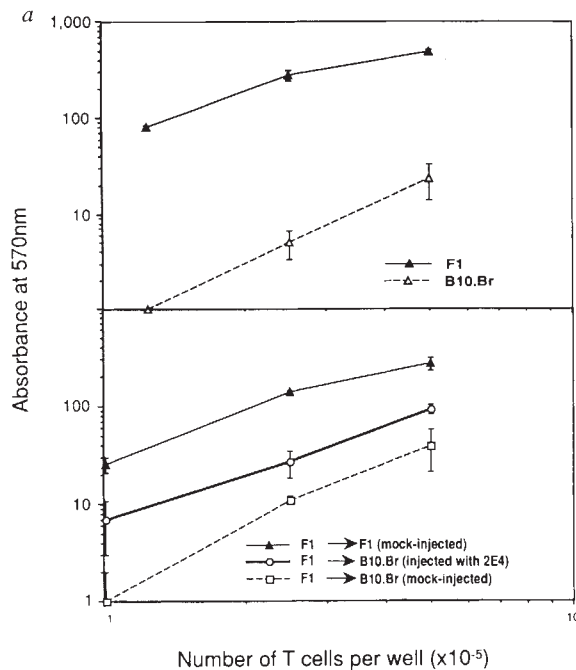
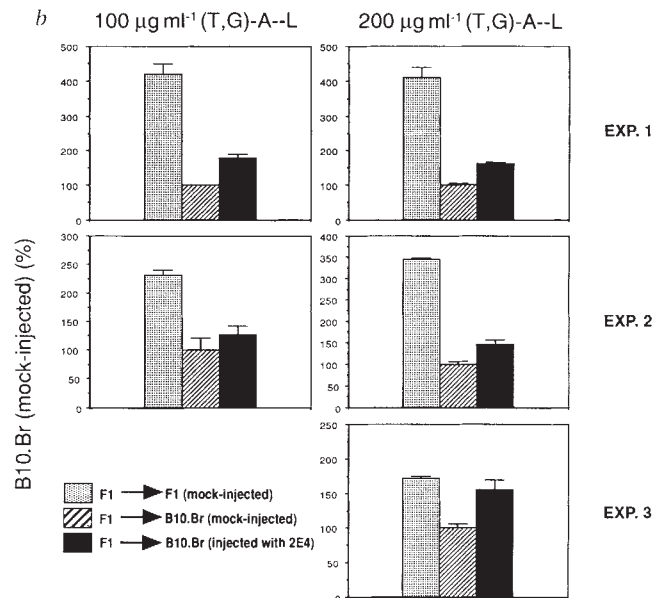
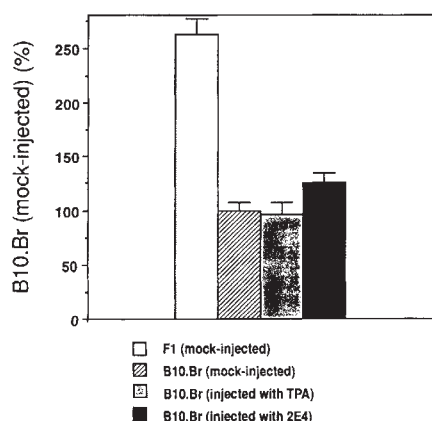


FIG. 3 Response to (T,G)-A--L by B10.BR mice reconstituted with BM cells and inoculated intrathymically with 2E4 cells. Response of F_1 mice to (T,G)-A--L was compared to that of B10.BR mice (a, upper panel). The response was also measured in $F_1 \rightarrow$ B10.BR chimaeras after intrathymic inoculation of 2E4 cells, or saline (mock-injected), and compared to $F_1 \rightarrow F_1$ chimaeras, mock-injected (a, lower panel, and b).

METHODS. B10.BR mice were γ -irradiated (9.5 Gy), and reconstituted i.v. with $10^7 F_1$ BM cells from donor mice treated one day earlier with 0.4 ml (i.p.) of rabbit anti-mouse T-lymphocyte serum (Accurate). Three weeks later, 10^7 2E4 cells were introduced in each thymic lobe²². After four weeks, mice were immunized with a subcutaneous injection of 50 μ g (T,G)-A--L (Miles Yeda) in complete Freund's adjuvant (Sigma). Seven days later, draining lymph nodes were collected from 3–4 mice per group and pooled; T cells

introduced 2E4 cells. These experiments show that intrathymic 2E4 cells could increase, probably by positive selection, the number of I-A^b-restricted T cells found in recipient animals.

Injection of epithelial cells may have increased the inefficient positive selection mediated by BM-derived cells by disrupting the thymic structure and allowing these cells access to the cortex and CD4⁺CD8⁺ thymocytes. Alternatively, a few I-A^b-restricted mature T cells could have been carried over in the BM inoculum, despite the anti-T lymphocyte serum treatment. To rule out these possibilities, we reconstituted B10.BR mice with fetal liver cells and compared the effect of the intrathymic injection of 2E4



enriched using nylon wool were cultured for 4 days in Click's medium, 1% normal mouse serum, with (T,G)-A--L and mitomycin C-treated splenocytes (5×10^5 per well). H-2^b splenocytes were added with T cells from control F_1 and reconstituted F_1 or B10.BR mice, whereas H-2^k splenocytes were used with T cells from control B10.BR mice. T-cell responses were measured using the MTT assay²³ and results are expressed as the mean of triplicates with antigen ($100 \mu\text{g ml}^{-1}$) minus background response in the absence of antigen (a). In b, similar titration curves from three independent experiments are plotted on a log versus log scale, and using the MKASSAY program (J.W.K.) a coefficient $\pm$ s.e. (horizontal offset of each curve) was obtained and expressed as the percentage of negative control ($F_1 \rightarrow$ B10.BR) chimaeras; mock-injected).

H-2^b cells to that of a thymic epithelial H-2^k cell line, TPA.B4. Injection of H-2^k cells did not cause an increase over background response but, as before, the response was higher to (T,G)-A--L with 2E4 cells (Fig. 4).

Overall, our data show that positive selection can be achieved by an epithelial cell line introduced into the thymus. Such cell lines can easily be manipulated and therefore be used to study the stromal molecules involved in positive selection.

Note added in proof: Since submission of this paper, Vukmanovic *et al.*²⁵ have reported positive selection of MHC class I using a similar approach. □

FIG. 4 Response to (T,G)-A--L by B10.BR mice reconstituted with fetal liver cells and inoculated intrathymically with 2E4 cells. The response was measured in B10.BR mice after reconstitution with fetal liver cells and intrathymic inoculation of saline alone (mock-injected), H-2^b 2E4 cells or H-2^k TPA.B4 epithelial cells. As a positive control the response of F_1 mice after reconstitution with F_1 fetal liver cells and intrathymic inoculation of saline (mock-injected) was also measured.

METHODS. Method as described in Fig. 3 legend, but here irradiated B10.BR mice were reconstituted i.v. with 6×10^6 liver cells from day 15.16 F_1 embryos. Three weeks later, 10^7 2E4 cells were introduced in each thymic lobe. As an additional negative control, a group of mice were injected with an epithelial cell line, TPA.B4, derived from a primary culture of AKR mouse thymus (gift from E. F. Potworowski) and immortalized with a chemical carcinogen²⁴. After 4 weeks, mice were immunized with (T,G)-A--L and 7 days later, nylon-wool-enriched lymph-node T cells were challenged *in vitro* with $100 \mu\text{g ml}^{-1}$ (T,G)-A--L for 4 days. Results are expressed as described for Fig. 3b. The key represents different host mouse groups, all reconstituted with fetal liver cells from F_1 embryos.

Received 13 July; accepted 29 October 1992.

1. Bevan, M. *Nature* **269**, 417–418 (1977).
2. Zinkernagel, R. M., Callahan, G. N., Cooper, S., Klein, P. A. & Klein, J. *J. exp. Med.* **147**, 882–896 (1978).
3. Benoist, C. & Mathis, D. *Semin. Immun.* **1**, 117–124 (1989).
4. Boyd, R. & Hugo, P. *Immun. Today* **12**, 71–79 (1991).
5. Pardoll, D. & Carrera, A. *Curr. Opin. Immun.* **4**, 162–165 (1992).
6. Berrih, S. *et al.* *J. Immun.* **135**, 1165–1171 (1985).
7. Ransom, J., Wu, R., Fischer, M., Spangrude, G. J. & Scollay, R. *Semin. Immun.* **134**, 180–190 (1991).
8. Bill, J. & Palmer, E. *Nature* **341**, 649–651 (1989).
9. Berg, L. J. *et al.* *Cell* **58**, 1035–1046 (1989).
10. Benoist, C. & Mathis, D. *Cell* **58**, 1027–1033 (1989).
11. Shortman, K., Egerton, M., Spangrude, G. J. & Scollay, R. *Semin. Immun.* **2**, 3–12 (1990).
12. Spangrude, G. J. & Scollay, R. *J. Immun.* **145**, 3661–3668 (1990).
13. Nishimura, T. *et al.* *J. Immun.* **145**, 4012–4017 (1990).
14. Lo, D. & Sprent, J. *J. Immun.* **137**, 1772–1775 (1986).
15. Montgomery, R. A. & Dalman, M. J. *J. Immun.* **147**, 554–560 (1991).
16. Fischer, M. *et al.* *J. Immun.* **146**, 3452–3456 (1991).

17. Bendelac, A. & Schwartz, R. H. *Nature* **353**, 68–71 (1991).
18. McDevitt, H. O. *et al.* *J. exp. Med.* **135**, 1259–1278 (1972).
19. Longo, D. L. & Davis, M. L. *J. Immun.* **130**, 2525–2527 (1983).
20. Bix, M. & Raulet, D. H. *Nature* **359**, 330–333 (1992).
21. Zlotnik, A., Shimonkevitz, R., Geffer, M., Kappler, J. & Marrack, P. *J. Immun.* **131**, 2814–2820 (1983).
22. Goldschneider, I., Komschlies, K. L. & Greiner, D. L. *J. exp. Med.* **163**, 1–17 (1986).
23. Mosmann, T. J. *Immun. Meth.* **65**, 55–63 (1983).
24. Potworowski, E. F., Turcotte, F., Beauchemin, C., Hugo, P. & Zelechowska, M. G. *In Vitro Cell Dev. Biol.* **22**, 557–560 (1986).
25. Vukmanovic, S., Granda, A. G., Faas, S. J., Knowles, B. B. & Bevan, M. J. *Nature* **359**, 729–732 (1992).

ACKNOWLEDGEMENTS. We thank G. Winslow and J. McCormack for critically reviewing the manuscript; E. F. Potworowski and M. Sela for the TPA.B4 cell line and (TG)-A–L respectively; H. Morales for technical assistance; and C. Gentry for direction of the animal care facility. This work was partially supported by the National Institutes of Health. P.H. is supported by a Centennial fellowship from the MRC of Canada.

Calcineurin mediates inhibition by FK506 and cyclosporin of recovery from α -factor arrest in yeast

Forrest Foor, Stephen A. Parent, Nancy Morin, Arlene M. Dahl, Naasa Ramadan, Gary Chrebet, Keith A. Bostian & Jennifer B. Nielsen*

Merck Research Laboratories, PO Box 2000, Bldg R80Y-2A97, Rahway, New Jersey 07065-0900, USA

* To whom correspondence should be addressed

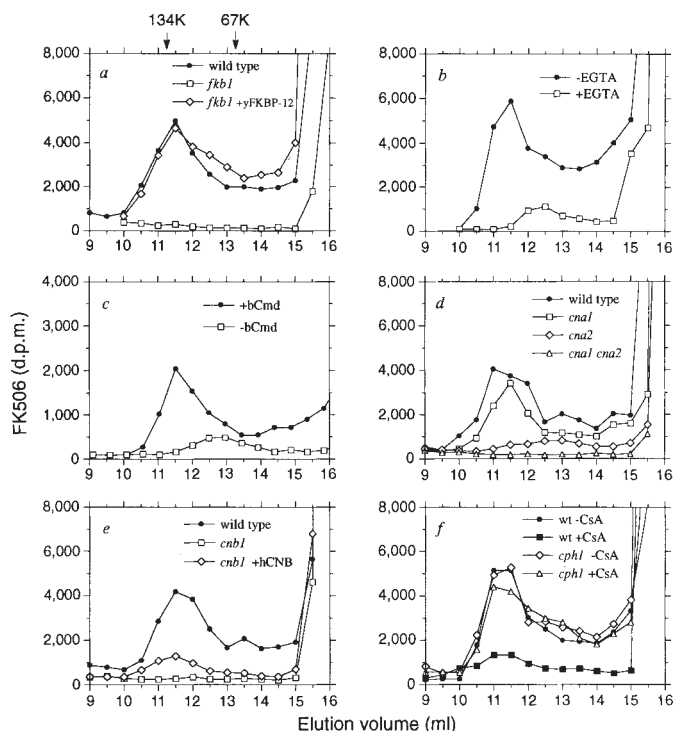
THE structurally unrelated immunosuppressants FK506 and cyclosporin A (CsA) act similarly, inhibiting a Ca^{2+} -dependent signal required for interleukin-2 transcription and T-cell activation¹. Each drug binds to its cytosolic receptor, FKBP-12 and cyclophilin,

respectively, and the drug–receptor complexes inhibit the Ca^{2+} /calmodulin-dependent protein phosphatase, calcineurin^{2–4}. In yeast, calcineurin has been implicated in recovery from α -mating factor arrest^{5,6}. Here we show that FK506 bound to yeast FKBP-12 appears to form a complex with yeast calcineurin. Moreover, recovery from mating factor arrest is highly sensitive to FK506 or CsA, and this sensitivity requires the presence of FKBP-12 or cyclophilin, respectively. These results define a key physiological target of an FK506- and CsA-sensitive signal pathway in yeast, suggest a high degree of mechanistic conservation with mammalian cells, and indicate that further examination of the yeast system should provide insight into the same process in T cells.

Saccharomyces cerevisiae contains at least two FK506-binding proteins, a cytosolic protein (yFKBP-12) of M_r 12,000 (12K) and a 13K membrane-associated protein (yFKBP-13)⁷. Here we report the characterization of a high-molecular-weight binding complex. Extracts incubated with [³H]dihydro-FK506 and fractionated on Superose-12 (Fig. 1; ref. 7) had a major peak of binding activity starting at an elution volume of 15 ml (partially

FIG. 1 FK506 forms a high-molecular-weight complex with yFKBP-12 and calcineurin. Superose-12 chromatography of labelled FK506 incubated with extracts from various yeast strains. All strains were derivatives of YFK005 (*MAT α*), YFK007 (*MAT α*) or their diploid YFK016 (ref. 10). Calcineurin mutants and wild-type (wt) strains were meiotic progeny of disrupted diploids. The *fkbl* mutant was YFK188 (ref. 10). In *a*, the complex was restored by addition of $1.2 \mu\text{g ml}^{-1}$ of recombinant yFKBP-12 (ref. 7) to the extract of the *fkbl* mutant. In *b*, formation of the complex was inhibited by addition of 2 mM EGTA to the sample and column buffer. In *c*, formation of the complex was greatly reduced in a calmodulin-depleted extract of YFK188 (see below) supplemented with $1.2 \mu\text{g ml}^{-1}$ of recombinant yFKBP-12, and was largely restored by addition of $8 \mu\text{g ml}^{-1}$ of bovine calmodulin (bCmd). In *e*, the complex was partially restored by addition of $20 \mu\text{g ml}^{-1}$ of human recombinant calcineurin B (hCNB) to the extract of the *cnb1* mutant. In *f*, formation of the complex was abolished by addition of $2 \mu\text{M}$ CsA to extracts of YFK007 (wt), but not to extracts of YFK547 (*cph1*).

METHODS. Mutant alleles were constructed as follows. DNA fragments corresponding to the published sequences (residues 1 to 2,689 for *CNA1* (ref. 8); 26 to 2,987 for *CNA2* (ref. 8); 31 to 680 for *CNB1* (ref. 9); and –522 to 823 for *CPH1* (ref. 12)) were synthesized by polymerase chain reaction using primers with a *Bss*HI site at the 5' end. The fragments were inserted into pCR1000 (Invitrogen). The *Bss*HI fragments of *CNA2* and *CNB1* were transferred to pBluescript II KS(+) (Stratagene). The *cna1* and *cna2* alleles were constructed essentially as described⁸. The *cnb1* allele was constructed by inserting the 4.8-kb *Hind*III–*Eco*RI *LYS2* fragment from Ylp600 (ref. 13) into the *StyI* site in the middle of the gene. The *cph1* allele was constructed by inserting the 3.8-kb *Bam*HI *ADE2* fragment from p669 (P. Linder, unpublished) between two *Bst*XI sites. Disruption fragments were excised with *Bss*HI, and the alleles introduced into diploid strains by one-step gene replacement. Loss of *CNA* gene products was confirmed by ¹²⁵I-labelled calmodulin binding on SDS–PAGE⁹, and disruption of *CNB1* was confirmed by western analysis with antisera to bovine calcineurin. Loss of *CPH1* was confirmed by LH-20 assay with ³H-CsA (ref. 14). Extracts were prepared as described⁷ from cells grown to early stationary phase. The calmodulin-depleted extract of YFK188 was prepared by chromatography of the crude extract on DEAE-cellulose as described⁶. Superose chromatography with



FK506 was essentially as described⁷, except that 100 mM Na_2SO_4 , 1 mM MgSO_4 and 0.1 mM CaCl_2 were included in all buffers. Binding is reported as d.p.m. per mg protein applied to the column, as measured in the crude extract.

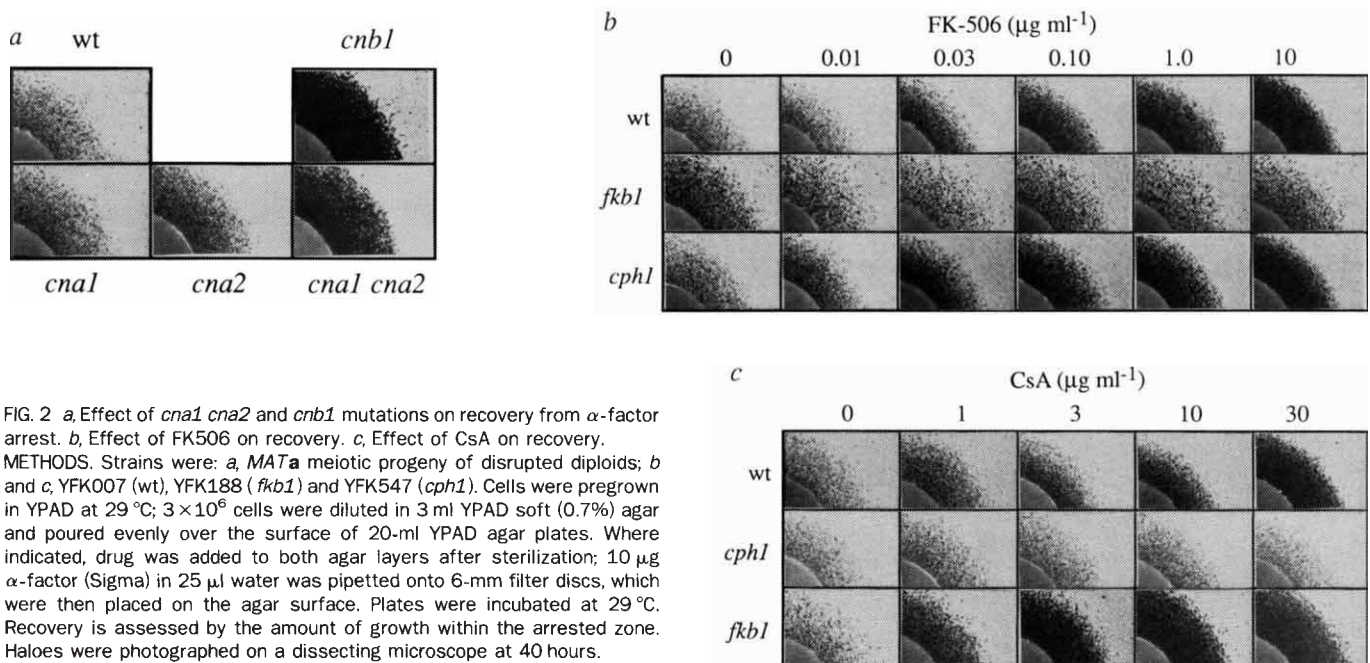


FIG. 2 *a*, Effect of *cna1 cna2* and *cnb1* mutations on recovery from α -factor arrest. *b*, Effect of FK506 on recovery. *c*, Effect of CsA on recovery. METHODS. Strains were: *a*, *MATa* meiotic progeny of disrupted diploids; *b* and *c*, YFK007 (wt), YFK188 (*fkb1*) and YFK547 (*cph1*). Cells were pregrown in YPAD at 29 °C; 3×10^6 cells were diluted in 3 ml YPAD soft (0.7%) agar and poured evenly over the surface of 20-ml YPAD agar plates. Where indicated, drug was added to both agar layers after sterilization; 10 μg α -factor (Sigma) in 25 μl water was pipetted onto 6-mm filter discs, which were then placed on the agar surface. Plates were incubated at 29 °C. Recovery is assessed by the amount of growth within the arrested zone. Haloes were photographed on a dissecting microscope at 40 hours.

shown), comprising FK506 bound to yFKBP-12 and yFKBP-13, and a minor peak of high- M_r material (100K complex), with $\sim 2\text{--}3\%$ of the total bound FK506, centred at 11.5 ml. yFKBP-12 is required as the 100K complex does not form in extracts of mutants (*fkb1*) lacking yFKBP-12 unless yFKBP-12 is added (Fig. 1a). yFKBP-13 is not a component, because the complex is unaltered in extracts of mutants (*fkb2*) lacking yFKBP-13 (not shown). Formation of the complex requires Ca^{2+} as it is largely eliminated by addition of EGTA (Fig. 1b), and calmodulin, because it is greatly reduced in extracts depleted of yeast calmodulin and can be reconstituted by addition of bovine calmodulin (Fig. 1c).

Because FK506 and mammalian FKBP-12 form a complex with mammalian calcineurin², it seemed likely that yeast calcineurin might be a component of the 100K complex. Two genes encoding yeast homologues of the catalytic subunit of calcineurin (*CNA1* and *CNA2*), and one encoding the regulatory subunit (*CNB1*), have been described^{5,6,8,9}. The 100K FK506 complex was reduced by 20% upon disruption of *CNA1* (Fig. 1d), and by 80% on disruption of *CNA2*. Disruption of both resulted in no detectable complex (Fig. 1d). Furthermore, disruption of *CNB1* completely abolished formation of the complex, and addition of human recombinant calcineurin B (which has 53% identity with the yeast protein) to the mutant extract partially restored it (Fig. 1e). Thus, yeast calcineurins are necessary for the formation of the complex. Although we have not strictly demonstrated that calcineurins are components of the complex, this seems likely on the basis of its size and by analogy with mammalian systems.

Recovery from G1 α -factor arrest in mating type *a* haploid cells (*MATa*) can be conveniently observed in a filter disk halo assay (Fig. 2). It has been reported that *cna1 cna2* and *cnb1* strains fail to resume growth in the continued presence of α -factor^{5,6}. We found that the *cna1* mutation had little effect on recovery, whereas the *cna2* mutation had a small, but visible, effect (Fig. 2a). Recovery was abolished in *cna1 cna2* or *cnb1* mutants. The influence of the *cna* and *cnb* mutations on recovery from α -factor arrest parallels their effects on the formation of the 100K complex.

Liu *et al.*² found that the complex of mammalian FKBP-12 with FK506 inhibits the protein phosphatase activity of mammalian calcineurin and suggested that this was the mechanism by which the drug exerts its immunosuppressive effect. The observation that overexpression of the catalytic A subunit, either

alone³ or in combination with the regulatory B subunit⁴, confers resistance to FK506 and CsA in Jurkat human T-cell lines, supports this model. We therefore investigated the effects of these drugs directly on recovery from mating factor arrest.

FK506 inhibited recovery from α -factor arrest strongly at 10 and 1 $\mu\text{g ml}^{-1}$, with significant effects seen at concentrations as low as 0.03 $\mu\text{g ml}^{-1}$ (Fig. 2b), which is more than three orders of magnitude less than that required to inhibit vegetative growth¹⁰. Recovery of mutants lacking yFKBP-12 (*fkb1*) was not affected by FK506, showing that yFKBP-12 is the receptor mediating the effect. As the *fkb1* mutation has little effect on sensitivity of vegetative cells to FK506 (ref. 10), growth sensitivity may be more relevant as a model of drug toxicity than immunosuppression.

The complex of mammalian cyclophilin with CsA also inhibits the protein phosphatase activity of mammalian calcineurin². Some inhibition of recovery from α -factor arrest was seen with CsA at 1 $\mu\text{g ml}^{-1}$ and increased at higher concentrations (Fig. 2c). Vegetative growth was unaffected at concentrations of up to 100 $\mu\text{g ml}^{-1}$ (not shown). Recovery of yeast mutants lacking the major cytosolic cyclophilin (*cph1*) was insensitive to the drug (Fig. 2c), indicating that *CPH1* is the primary mediator of this effect, as well as of the growth inhibition seen in CsA-hypersensitive mutants¹¹. Although a stable complex of labelled CsA with its binding protein and yeast calcineurin was not detected by Superose-12 chromatography, formation of the 100K FK506 complex was largely eliminated by addition of excess CsA (Fig. 1f). This competition was absent in the *cph1* mutant, suggesting that CsA, when bound to cyclophilin, is able to bind calcineurin and displace FK506/yFKBP-12.

FK506 and CsA inhibit recovery from α -mating factor arrest in yeast through interactions strikingly similar to those involved in the inhibition of T-cell activation. This inhibition occurs at drug concentrations much lower than those affecting vegetative growth, the only drug-sensitive property previously reported in lower eukaryotes. The high level of conservation of these drug-receptor calcineurin interactions suggests that other elements of the Ca^{2+} /calcineurin-dependent signal transduction pathway may also be conserved, offering new possibilities for the investigation of these signal pathways in yeast and T cells. □

Received 24 July; accepted 26 October 1992.

1. Sigal, N. H. & Dumont, F. J. A. *Rev. Immun.* 519–560 (1992).
2. Liu, J. *et al. Cell* 66, 807–815 (1991).

3. O'Keefe, S. J. *et al. Nature* **357**, 692-694 (1992).
4. Clipstone, N. A. & Crabtree, G. R. *Nature* **357**, 695-697 (1992).
5. Cyert, M. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 7376-7380 (1991).
6. Cyert, M. S. & Thorner, J. *Molec. cell. Biol.* **12**, 3460-3469 (1992).
7. Nielsen, J. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 7471-7475 (1992).
8. Liu, Y. *et al. Molec. gen. Genet.* **227**, 52-59 (1991).
9. Kuno, T. *et al. Biochem. biophys. Res. Commun.* **180**, 1159-1163 (1991).
10. Brizuela, L. *et al. Molec. cell. Biol.* **11**, 4616-4626 (1991).
11. Tropea, M. *et al. Nature* **342**, 953-955 (1989).
12. Haendler, B. *et al. Gene* **83**, 39-46 (1989).
13. Barnes, D. A. & Thorner, J. *Molec. cell. Biol.* **6**, 2828-2838 (1986).
14. Handschumacher, R. E. *et al. Science* **226**, 544-547 (1984).

ACKNOWLEDGEMENTS. We thank G. Salituro for FK506, P. Durette for CsA, A. Marcy for the *E. coli* *FKB1* expression plasmid, J. Burbaum for human CNB, D. Wolff for antisera to bovine calcineurin, and N. Sigal, J. Siekierka, G. Wiederrecht, E. O'Neill, S. O'Keefe and M. Tocci for discussion and critically reading the manuscript.

Isozyme-selective stimulation of phospholipase C- β 2 by G protein $\beta\gamma$ -subunits

Montserrat Camps*, Amanda Carozzi†, Petra Schnabel*, Alexander Scheer*, Peter J. Parker† & Peter Gierschik*‡

* Molecular Pharmacology Division, German Cancer Research Center, 6900 Heidelberg, Germany

† Protein Phosphorylation Laboratory, Imperial Cancer Research Fund, London WC2A 3PX, UK

HYDROLYSIS by phospholipase C (PLC) of phosphatidylinositol 4,5-bisphosphate is a key mechanism by which many extracellular signalling molecules regulate functions of their target cells^{1,2}. At least eight distinct isoforms of PLC are recognized in mammalian cells^{3,4}. Receptor-controlled PLC is often regulated by G proteins, which can be modified by pertussis toxin in some cells but not in others^{5,6}. In the latter cells, PLC- β 1, but not PLC- γ 1 or PLC- δ 1, may be activated by members of the α_q -subfamily of the G protein α -subunits⁷⁻¹⁰. An unidentified PLC in soluble fractions of cultured human HL-60 granulocytes is specifically stimulated by G protein $\beta\gamma$ subunits purified from retina and brain¹¹. Identification of a second PLC- β complementary DNA (PLC- β 2) in an HL-60 cell cDNA library⁹ prompted us to investigate the effect of purified G protein $\beta\gamma$ subunits on the activities of PLC- β 1 and PLC- β 2 transiently expressed in cultured mammalian cells. We report here that PLC- β 1 and PLC- β 2 were stimulated by free $\beta\gamma$ subunits and that PLC- β 2 was the most sensitive to $\beta\gamma$ stimulation. Thus stimulation of PLC by $\beta\gamma$ subunits is isozyme-selective and PLC- β 2 is a prime target of $\beta\gamma$ stimulation. Activation of PLC- β 2 by $\beta\gamma$ subunits may be an important mechanism by which pertussis toxin-sensitive G proteins stimulate PLC.

Transfection of COS-1 cells with the mammalian expression vector pMT2 containing the cDNAs of PLC- β 1 and PLC- β 2 led to a marked increase in inositol phosphate formation by the corresponding cell lysates relative to the formation observed for cells transfected with pMT2 alone (Fig. 1a). Inositol phosphate formation by lysates of cells transfected with pMT2-PLC- β 1 or pMT2-PLC- β 2 was absolutely dependent on and greatly stimulated by free Ca^{2+} . The Ca^{2+} sensitivity was similar for both lysates. The immunochemical analysis of lysates from transfected cells using monospecific antisera reactive against PLC- β 1 and PLC- β 2 revealed that the two PLC isoforms were synthesized in significant quantities as native polypeptides with apparent M_r s of ~165K and 155K, respectively (Fig. 1b). The lysates also contained several immunoreactive polypeptides with smaller M_r s, which presumably correspond to proteolytic degradation products of the two native PLCs.

The effect of purified G protein $\beta\gamma$ subunits on the activities

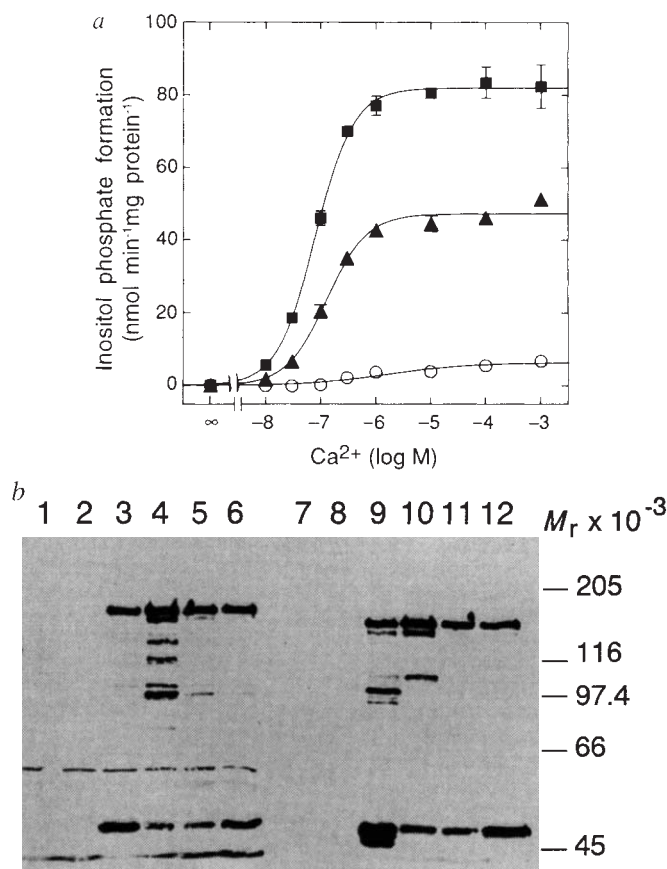


FIG. 1 Expression of PLC- β 1 and PLC- β 2 in COS-1 cells. **a**, Effect of Ca^{2+} on inositol phosphate formation by extracts of COS-1 cells. Cells were transfected with the vector pMT2 without insert (○), with pMT2-PLC- β 1 (▲) or pMT2-PLC- β 2 (■). Cell extracts (0.3 μg protein per assay) were incubated for 30 min with phospholipid vesicles containing $\text{PtdIns}(4,5)\text{P}_2$ at the concentrations of free Ca^{2+} indicated at the abscissa. **b**, Immunochemical analysis of PLC- β 1 and PLC- β 2 expressed in COS-1 cells. COS-1 cells were transfected with vector without insert (lanes 2 and 8), or vector containing the cDNAs of PLC- β 1 (lanes 3-6) or PLC- β 2 (lanes 9-12). Extracts of non-transfected cells (lanes 1 and 7) and of transfected cells (lanes 2-6 and 8-12) were subjected to SDS-PAGE (150 μg protein per lane) and immunoblotting was done using antisera reactive against PLC- β 1 (lanes 1-6) or PLC- β 2 (lanes 7-12). The two major ~165K and ~155K immunoreactive proteins detected by the two sera in lanes 3-6 and 9-12 comigrated on immunoblots with PLC- β 1 and PLC- β 2 present in bovine brain and HL-60 cell membranes, respectively (results not shown). **METHODS.** The cDNAs encoding bovine PLC- β 1 (ref. 22) and human PLC- β 2 (ref. 9) were ligated into the mammalian expression vector pMT2 (ref. 23). COS-1 cells were propagated in DMEM supplemented with 10% FCS. Equal amounts of plasmid DNA (15 μg per 10-cm dish) were transfected into COS-1 cells by lipofection using transfectam (Promega) as recommended by the manufacturer. 48 h after transfection, cells were rinsed 3 times with PBS followed by addition of ice-cold extraction buffer A (500 μl per 10-cm dish) containing 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 10 mM EGTA, 37 mM sodium cholate, 3 mM benzimidazole, 43 mM β -mercaptoethanol and 100 μM phenylmethylsulphonyl fluoride (PMSF). Cells were scraped into extraction buffer, incubated for 1 h at 4°C with gentle vortexing every 10 min, and then centrifuged at 18,000g for 20 min. Supernatants were diluted ~200-fold with buffer A containing no EDTA and 18.5 mM sodium cholate and assayed for PLC activity as described¹², except that EGTA (3 mM) was used as a chelating agent. The final concentrations of sodium cholate and sodium deoxycholate were 1.3 mM and 3.4 mM, respectively. The formation of inositol phosphates was determined as described²⁴. SDS-PAGE and immunoblotting was done as described²⁴, except that immunoreactive proteins were visualized using the ECL western blotting detection system (Amersham). Antisera reactive against PLC- β 1 and PLC- β 2 were prepared by coupling synthetic peptides representing the carboxy-terminal regions of PLC- β 1 (NH_2 -GENPGKEFDTP- COOH ; single-letter amino-acid code) or PLC- β 2 (NH_2 -QDPLIAKADAQESRL- COOH) to keyhole limpet haemocyanin and immunizing rabbits.

‡ To whom correspondence should be addressed.

TABLE 1 Effect of $\beta\gamma_i$ on inositol phosphate formation by extracts of COS-1 transfectants

Transfection	$\beta\gamma_t$	$\text{Ca}^{2+}/\text{DOC}$	Inositol phosphate formation			
			Control	pMT2	pMT2-PLC- β 1	pMT2-PLC- β 2
			nmol min ⁻¹ mg protein ⁻¹			
1	—	+	ND	1.3±1.0	1.4±1.0	3.0±0.8
	+	+	ND	1.0±0.9	3.1±0.8	39.0±2.7
	—	++	ND	3.8±0.8	60.0±4.4	97.4±8.3
2	—	+	0	0	1.8±1.1	3.8±1.2
	+	+	0	0	7.4±1.0	25.0±2.0
	—	++	1.5±0.9	1.4±1.3	50.8±6.2	41.0±1.8
3	—	+	0	0	1.4±0.6	2.3±0.5
	+	+	1.3±0.8	0	6.7±1.3	12.0±0.5
	—	++	1.7±0.5	0.9±0.5	32.0±1.1	18.3±1.7
4	—	+	ND	0.9±0.8	1.9±0.8	2.8±0.9
	+	+	ND	1.5±0.8	4.5±1.0	18.7±2.5
	—	++	ND	4.8±1.3	35.8±2.3	44.5±0.8

Inositol phosphate formation by extracts of COS-1 transfectants. Cell extracts (0.3 μg protein per assay) from non-transfected cells (control) or cells transfected with the indicated plasmid DNAs were incubated in the absence or presence of $\beta\gamma_i$ (2 μM) with phospholipid vesicles containing PtdInsP_2 . The assay was done for 30 min in the presence of either 0.1 μM free Ca^{2+} and 0.9 mM sodium deoxycholate (+), or 10 μM free Ca^{2+} and 3.4 mM sodium deoxycholate (++). Each value corresponds to the mean $\pm$ s.d. of triplicate determinations. DOC, sodium deoxycholate; ND, not determined.

of PLC- β 1 and PLC- β 2 expressed in COS-1 cells was determined using hydrophilic, detergent-free $\beta\gamma$ subunits of retinal transducin ($\beta\gamma_i$) (Table 1). $\beta\gamma_i$ had no effect on the inositol phosphate formation by lysates of non- or pMT2-transfected cells. In contrast, $\beta\gamma_i$ clearly stimulated the formation of inositol phosphates by lysates of cells transfected with pMT2-PLC- β 1 and pMT2-PLC- β 2. Most interestingly, however, the degree of stimulation by $\beta\gamma_i$ was considerably higher for the cells transfected with pMT2-PLC- β 2 than for those transfected with pMT2-PLC- β 1. A summary of the results accounting for the differences in expressed PLC activities is shown in Fig. 2a; this shows that the net increase in inositol phosphate formation induced by $\beta\gamma_i$ was about fivefold higher for cells expressing PLC- β 2 than for cells expressing PLC- β 1.

The dependence of inositol phosphate formation by lysates of cells expressing PLC- β 1 or PLC- β 2 on the concentration of $\beta\gamma_i$ is shown in Fig. 2b. $\beta\gamma_i$ did not affect the inositol phosphate formation by extracts of pMT2-transfected cells. In contrast, $\beta\gamma_i$ led to an increase in the activity of extracts from pMT2-PLC- β 2-transfected cells, which was ~ 20 -fold at 3 μM $\beta\gamma_i$. In lysates from cells expressing PLC- β 1, the stimulation by this concentra-

tion of $\beta\gamma_i$ was only about threefold. Saturation of the effect of $\beta\gamma_i$ on inositol phosphate formation by PLC- β 1 or PLC- β 2 was not achieved even at higher concentrations of $\beta\gamma_i$. To examine the specificity of the stimulation of PLC- β 2 by $\beta\gamma_i$, we determined the effect of GDP-liganded α_i ($\alpha_{i,\text{GDP}}$) on $\beta\gamma_i$ -stimulated phosphatidylinositol 4,5-bisphosphate hydrolysis by PLC- β 2 (Fig. 3). $\alpha_{i,\text{GDP}}$ completely eliminated the stimulatory effect of $\beta\gamma_i$. In contrast, $\alpha_{i,\text{GDP}}$ had no effect on inositol phosphate formation by lysates of PLC- β 2-transfected cells when tested at the same concentrations in the absence of $\beta\gamma_i$ (results not shown). Thus, PLC- β 2 is specifically stimulated by free $\beta\gamma$, but not by the $\alpha\beta\gamma$ heterotrimer.

These results demonstrate that both PLC- β 1 and PLC- β 2 are stimulated by G protein $\beta\gamma$ subunits and that PLC- β 2 is a major target of $\beta\gamma$ -subunit stimulation. These findings, together with our previous observation of an unidentified $\beta\gamma_i$ -resistant PLC in HL-60 granulocytes¹¹ and our unpublished observation that purified PLC- δ 1 showed no response to $\beta\gamma_i$, strongly suggest that the stimulation of PLC- β 2 by $\beta\gamma_i$ is specific for members of the PLC- β subfamily and is isozyme-selective within this group of PLCs. Clearly, PLC- β 1 is activated by all four members

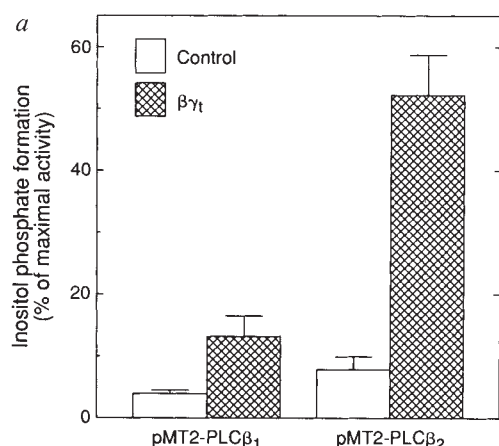
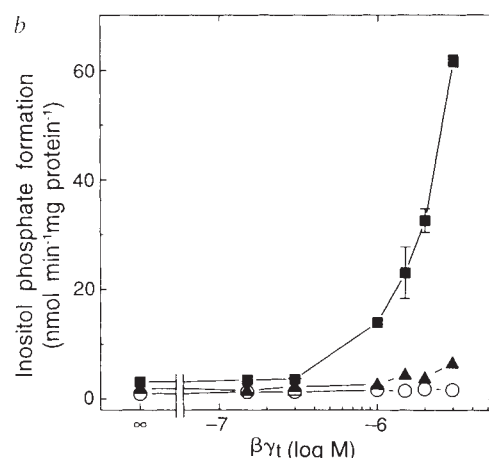


FIG. 2 a, Effect of $\beta\gamma_i$ on inositol phosphate formation by extracts of COS-1 cells expressing PLC- β 1 or PLC- β 2. The inositol phosphate formation was determined at 0.1 μM free Ca^{2+} and 0.9 mM sodium deoxycholate and is given in per cent of the maximal activity determined at 10 μM free Ca^{2+} and 3.4 mM sodium deoxycholate. The values are derived from the results given in Table 1 and correspond to the means $\pm$ s.e.m. ($n=4$). b, Effect of $\beta\gamma_i$ on inositol phosphate formation by extracts of COS-1 transfectants.



COS-1 cells were transfected with pMT2 (○), pMT2-PLC- β 1 (▲), or pMT2-PLC- β 2 (■). Cell extracts (0.3 μg protein per assay) were incubated for 40 min with phospholipid vesicles containing $\text{PtdIns}(4,5)\text{P}_2$ at the concentrations of $\beta\gamma_i$ indicated at the abscissa. The reaction mixture contained 0.1 μM free Ca^{2+} and 0.9 mM sodium deoxycholate. METHODS. The purification of $\beta\gamma_i$ and the conditions used to assay the effect of $\beta\gamma_i$ on PLC activity have been described¹¹.

of the α_q family (α_q , α_{11} , α_{14} , and α_{16})⁷⁻¹⁰. PLC- $\beta 2$ seems to be resistant to stimulation by α_q , α_{11} , and α_{14} , but sensitive to stimulation by α_{16} (refs 9, 10). Our finding that both PLC- $\beta 1$ and PLC- $\beta 2$ are stimulated by $\beta\gamma$ subunits strongly suggests that these PLCs, similar to the type II and IV adenylyl cyclases¹²⁻¹⁴, are regulated by both G protein α - and $\beta\gamma$ subunits.

An important issue is the relevance of $\beta\gamma$ -mediated stimulation of inositol phosphate formation to receptor-mediated transmembrane signalling in native plasma membranes and in intact cells. Clearly, this signalling pathway is much more likely to occur in cells expressing predominantly PLC- $\beta 2$ rather than PLC- $\beta 1$. Membranes of peripheral blood neutrophils and HL-60 granulocytes contain significant quantities of PLC- $\beta 2$, but lack PLC- $\beta 1$ (our unpublished results). The results shown in Fig. 3 indicate that relatively high concentrations of $\beta\gamma$ subunits are required to stimulate inositol phosphate formation by PLC- $\beta 2$. But similarly high concentrations of $\beta\gamma_i$ were required to stimulate retinal phospholipase A_2 (ref. 15). Differences have been noted in the relative ability of various $\beta\gamma$ dimers to stimulate atrial phospholipase A_2 (ref. 16) or adenylyl cyclase¹². Will different species of $\beta\gamma$ dimers be more potent stimulators of PLC than $\beta\gamma_i$? In granulocyte plasma membranes, the calculated concentration of $\beta\gamma$ subunits is $\sim 70 \mu\text{M}$ ¹¹. These membranes also contain high numbers of receptors for *N*-formyl methionine-containing chemotactic peptides, which are coupled to stimulation of PLC through the pertussis toxin-sensitive G proteins G_{12} and/or G_{13} (ref. 17). G_{12} and G_{13} are highly abundant in granulocyte membranes and much of the cellular $\beta\gamma$ subunits is probably complexed to α_{12} and α_{13} . Agonist occupancy of the formyl peptide receptor may lead to a transient, pertussis toxin-sensitive increase in the concentration of free $\beta\gamma$ subunits which is sufficiently high to stimulate PLC- $\beta 2$. Thus, in granulocytes, stimulation of PLC- $\beta 2$ by the free $\beta\gamma$ subunit is likely to be the mechanism by which the pertussis toxin-sensitive G proteins G_{12} and/or G_{13} stimulate PLC.

If this hypothesis is correct, one would predict that other G-protein-coupled receptors capable of inducing the release of a similarly high concentration of $\beta\gamma$ subunits are capable of stimulating inositol phosphate formation. Recent results obtained by expressing a variety of G-protein-coupled receptors in cultured mammalian cells suggest that this may be the case. For example, several receptors well known to inhibit adenylyl cyclase through G_i (for example, the m2 and m4 muscarinic

receptors, the 5-HT_{1A} receptor, α_2 -adrenergic and D2 dopaminergic receptors¹¹) and several adenylyl cyclase stimulatory receptors (for example, the TSH¹⁸, LH¹⁹, calcitonin²⁰ and PTH receptors²¹) also stimulate PLC. Even receptors not usually coupled to phospholipase C may thus interact through G protein $\beta\gamma$ subunits with this effector enzyme under certain circumstances, for example in cells expressing PLCs highly sensitive to $\beta\gamma$ stimulation such as PLC- $\beta 2$.

These results indicate that the degree of cross-talk between G proteins and effectors and thus the flexibility of G-protein-regulated transmembrane signalling is much higher than thought. Even a single G protein may regulate more than one effector by virtue of its subunit structure and that more than one G protein, such as G_{12} and G_{13} , may regulate the same effector (PLC- $\beta 2$) by releasing common subunits, the $\beta\gamma$ dimers. □

Received 23 July; accepted 9 October 1992.

- Berridge, M. J. & Irvine, R. F. *Nature* **341**, 197-205 (1989).
- Nishizuka, Y. *Nature* **334**, 661-665 (1988).
- Meldrum, E., Parker, P. J. & Carozzi, A. *Biochim. biophys. Acta* **1092**, 49-71 (1991).
- Rhee, S. G. & Choi, K. D. *Adv. Cyclic Nucleotide Prot. Phosph. Res.* **26**, 35-61 (1992).
- De Vivo, M. & Gershengorn, M. C. in *ADP-Ribosylating Toxins and G Proteins: Insights into Signal Transduction* (eds Moss, J. & Vaughan, M.) 267-293 (American Society for Microbiology, Washington DC, 1990).
- Birnbaumer, L., Abramowitz, J. & Brown, A. M. *Biochim. biophys. Acta* **1031**, 163-224 (1990).
- Taylor, S. J., Chae, H. Z., Rhee, S. G. & Exton, J. H. *Nature* **350**, 516-518 (1991).
- Smrcka, A. V., Hepler, J. R., Brown, K. O. & Sternweis, P. C. *Science* **251**, 804-807 (1991).
- Park, D. J., John, D.-Y., Kriz, R. W., Knopf, J. L. & Rhee, S. G. *J. Biol. Chem.* **267**, 16048-16055 (1992).
- Lee, C. H., Park, D. J., Wu, D., Rhee, S. G. & Simon, M. I. *J. Biol. Chem.* **267**, 16044-16047 (1992).
- Camps, M. et al. *Eur. J. Biochem.* **206**, 821-831 (1992).
- Tang, W.-J. & Gilman, A. G. *Science* **254**, 1500-1503 (1991).
- Gao, B. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **88**, 10178-10182 (1991).
- Federman, A. D., Conklin, B. R., Schrader, K. A., Reed, R. R. & Bourne, H. R. *Nature* **356**, 159-161 (1992).
- Jelsema, C. L. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3623-3627 (1987).
- Schmidt, C. J. & Neer, E. J. *J. Biol. Chem.* **266**, 4538-4533 (1991).
- Gierschik, P., Sidiropoulos, D. & Jakobs, K. H. *J. Biol. Chem.* **264**, 21470-21473 (1989).
- Van Sande, J. et al. *Molec. cell. Endocrin.* **74**, R1-R6 (1990).
- Guderman, T., Birnbaumer, M. & Birnbaumer, L. *J. Biol. Chem.* **267**, 4479-4488 (1992).
- Chabre, O. et al. *Molec. Endocrin.* **6**, 551-556 (1992).
- Abou-Samra, A.-B. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2732-2736 (1992).
- Katan, M. et al. *Cell* **54**, 171-177 (1988).
- Kaufman, R. J. *Meth. Enzym.* **185**, 487-511 (1990).
- Gierschik, P., Codina, J., Simons, C., Birnbaumer, L. & Spiegel, A. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 727-731 (1985).

ACKNOWLEDGEMENTS. We thank R. W. Kriz for the PLC- $\beta 2$ cDNA, S. G. Rhee for ref. 9 before publication and E. Strohmaier for technical assistance. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the Fritz Thyssen Stiftung. M.C. received a grant from the Commission of the European Communities.

Subunits $\beta\gamma$ of heterotrimeric G protein activate $\beta 2$ isoform of phospholipase C

Arieh Katz, Dianqing Wu & Melvin I. Simon*

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

THE activation of heterotrimeric G proteins results in the exchange of GDP bound to the α -subunit for GTP and the subsequent dissociation of a complex of the β - and γ -subunits ($G_{\beta\gamma}$). The α -subunits of different G proteins interact with a variety of effectors¹⁻⁷, but less is known about the function of the free $G_{\beta\gamma}$ complex. $G_{\beta\gamma}$ has been implicated in the activation of a cardiac potassium channel⁸, a retinal phospholipase A_2 (ref. 9) and a specific receptor kinase¹⁰, and *in vitro* reconstitution experiments indicate that the $G_{\beta\gamma}$ complex can act with G_α subunit to modulate the activity of different isoforms of adenylyl cyclase¹¹. Of two phospholipase activities that can be separated in extracts of HL-60 cells, purified $G_{\beta\gamma}$ is found to activate one of them¹². Here we report that in co-transfection assays $G_{\beta\gamma}$ subunits specifically

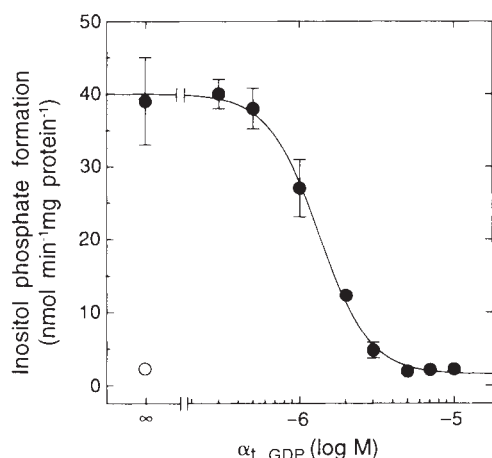


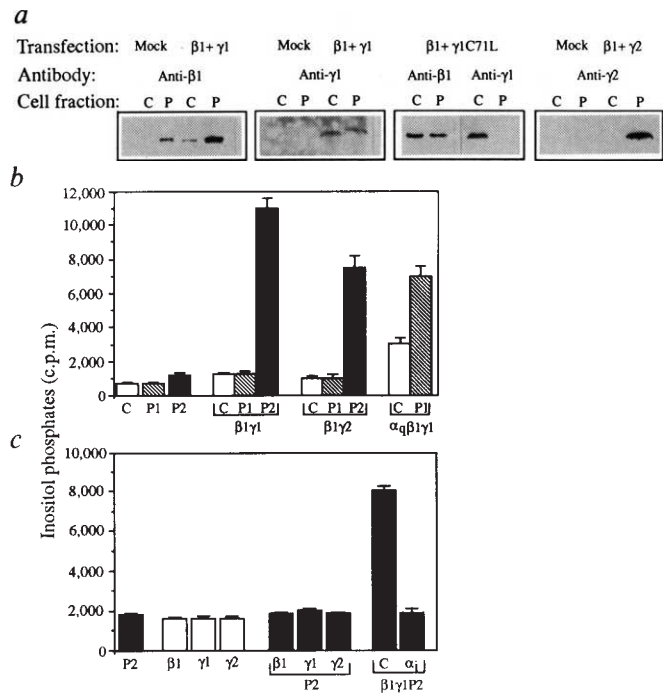
FIG. 3 Reversal of the $\beta\gamma_i$ -dependent stimulation of PLC- $\beta 2$ by GDP-liganded α_t . COS-1 cells were transfected with pMT2-PLC- $\beta 2$. Aliquots of the cell extracts (0.3 μg protein per assay) were incubated for 40 min in the absence of $\beta\gamma_i$ (○) or in the presence of $\beta\gamma_i$ (2 μM) and increasing concentrations of GDP-liganded α_t (●) with phospholipid vesicles containing PtdIns(4, 5) P_2 . The reaction mixture contained 0.1 μM free Ca^{2+} and 0.9 mM sodium deoxycholate.

METHODS. The purification of α_t and the conditions used to assay the effect of α_t on $\beta\gamma_i$ -stimulated PLC activity have been described¹¹.

* To whom correspondence should be addressed.

FIG. 1 Effect of $G_{\beta_1\gamma_1}$ and $G_{\beta_1\gamma_2}$ subunits on PLC- β 1 and PLC- β 2 isoenzymes in transiently transfected COS-7 cells. **a**, Immunoblot analysis of the expression of cDNA clones corresponding to β_1 , γ_1 , γ_1 C71L and γ_2 subunits in the cytosol (C) and particulate (P) fractions of transfected COS-7 cells. Cells (1×10^5) transfected as indicated, were fractionated and analysed. Mock-transfected cells are cells transfected with vector pcDNA-I (Invitrogen). Only sections of the blots corresponding to the location of β - and γ -subunits are shown. **b** and **c**, Accumulation of inositol phosphates in COS-7 cells transfected with expression vectors carrying cDNAs corresponding to: β -galactosidase (C), PLC- β 1 (P1), PLC- β 2 (P2), G_{β_1} , G_{γ_1} , G_{γ_2} , G_{α_q} and $G_{\alpha_{12}}$ (α_i). The cDNAs transfected are indicated beneath each column. Data represent duplicate determinations in a single experiment, and error bar gives the range of duplicates; three other experiments gave similar results. The relative level of each expressed component was similar in the various transfections as determined by western blotting.

METHODS. COS-7 cells were maintained in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal calf serum (FCS). 1×10^5 cells per well were seeded in 12-well plates a day before transfection. All cDNAs were inserted into cytomegalovirus (CMV) expression vectors pCDM8.1 (ref. 28), pCIS (ref. 29) and pcDNA-I as follows: β_1 -pCDM8.1, γ_1 -pcDNA-I, γ_1 C71L-pcDNA-I, γ_2 -pCDM8.1, PLC- β 1-pCIS, PLC- β 2-pCIS, G_{α_q} -pCIS and $G_{\alpha_{12}}$ -pCIS. The total amount of DNA in all transfections was $1 \mu\text{g}$ per well. The amount of each type of DNA in each set of experiments was equal and pCIS encoding β -galactosidase was used to maintain the amount of DNA constant. The efficiency of transfection was consistent (30–40%) as determined by transfecting the cells with β -galactosidase-pCIS, and 2 d later staining them with 5-bromo-4-chloro-3-indolyl- β -D-galactoside. DNA ($1 \mu\text{g}$) was mixed with $5 \mu\text{l}$ lipofectin (BRL) in 0.5 ml Opti-MEM (BRL) and added to the cells; 5 h later 0.5 ml 20% FCS in DMEM was added to each well. For immunoblot analysis, cells were collected 48 h post-transfection and lysed in hypotonic buffer (50 mM HEPES, 0.2 mM EGTA, pH 7.0, 0.01% soybean trypsin inhibitor, 1 mM PMSF, $0.5 \mu\text{g ml}^{-1}$ leupeptin, $0.5 \mu\text{g ml}^{-1}$ pepstatin A) by freezing and thawing. The suspension was then spun and the cytosolic fraction (supernatant) separated from the particulate fraction. Samples were resolved by SDS-PAGE (10% and 15% acrylamide for β_1 and γ -subunits respectively) and electroblotted. Blots were incubated with antipeptide antibodies directed against the NH_2 -terminal portion of β_1 , γ_1 and γ_2 subunits as before¹⁸ and processed using an enhanced chemiluminescence detection system (Amersham). To quantify inositol phosphates in transfected cells, cells were washed with phosphate-buffered saline (PBS) 24 h post-transfection and incubated in 0.4 ml medium of inositol-free DMEM with 10% dialysed



FCS containing $10 \mu\text{Ci ml}^{-1}$ of [^3H]myo-inositol (Amersham). Cells were washed 24 h later with PBS and 200 μl was added of inositol-free medium containing 10 mM LiCl with or without ligand (10 μM carbachol). After incubation for 25 min at 37°C 200 μl cold 10% perchloric acid and 20 μl phytic acid (20 mg ml^{-1}) were added and the cells were incubated on ice for 10 min. Then 200 μl supernatant was transferred to a microcentrifuge tube and neutralized with 2 M KOH. After centrifugation, the supernatant was loaded on a 0.5-ml AG1-X8 anion exchange column (200–400 mesh, formate form; BioRad). Separation was essentially as described³⁰; 0.5 ml of the elution was mixed with 10 ml scintillation cocktail (BCS; Amersham) and counted.

activate the β 2 and not the β 1 isoform of phospholipase, which acts on phosphatidylinositol. We use transfection assays to show also that receptor-mediated release of $G_{\beta\gamma}$ from G proteins that are sensitive to pertussis toxin can result in activation of the phospholipase. This effect may be the basis of the pertussis-toxin-sensitive phospholipase C activation seen in some cell systems (reviewed in refs 13 and 14).

G-protein signalling pathways have been reconstituted by transient transfection into COS-7 cells^{5,15–17}. We have previously used co-transfection of the complementary DNAs encoding phospholipase C and G protein α -subunits to demonstrate the specific interaction of α -subunits of the G_{α_q} subfamily with the β 1 isoform of phospholipase C (ref. 5). A similar system was used here to test the specificity of interaction of the $G_{\beta\gamma}$ subunits with phospholipase C. Cells were co-transfected with expression vectors carrying various cDNA inserts corresponding to combinations of G protein β - and γ -subunits. Figure 1a shows an immunoblot analysis of the particulate and cytosol fractions of co-transfected COS-7 cells. As previously shown^{18–20}, $G_{\beta\gamma}$ complexes are, for the most part, associated with the membrane fraction, but some of the expressed $G_{\beta_1\gamma_1}$ is found in the cytosol. A mutant of the G_{γ_1} subunit (leucine substituted for cysteine at position 71: γ_1 C71L) that cannot be modified by isoprenylation is confined to the cytosol and shifts the localization of the co-transfected G_{β_1} subunit to the cytosol as previously reported^{18–20}. Antisera detect low levels of endogenous G_{β_1} subunits but no endogenous G_{γ_1} or G_{γ_2} protein.

Figure 1b shows the results of co-transfection of cDNAs encoding the two isoforms of phospholipase C- β (PLC- β) together with $G_{\beta\gamma}$ cDNAs. There is little or no effect of transfection with PLC- β 1 or PLC- β 2 alone, but co-transfection with

either $G_{\beta_1\gamma_1}$ or $G_{\beta_1\gamma_2}$ leads to a 5–10-fold stimulation of PLC activity in the PLC- β 2-transfected cells and not in the PLC- β 1-transfected cells. In the control experiment (Fig. 1b), co-transfection of cells with G_{α_q} cDNA together with $G_{\beta_1\gamma_1}$ and PLC- β 1

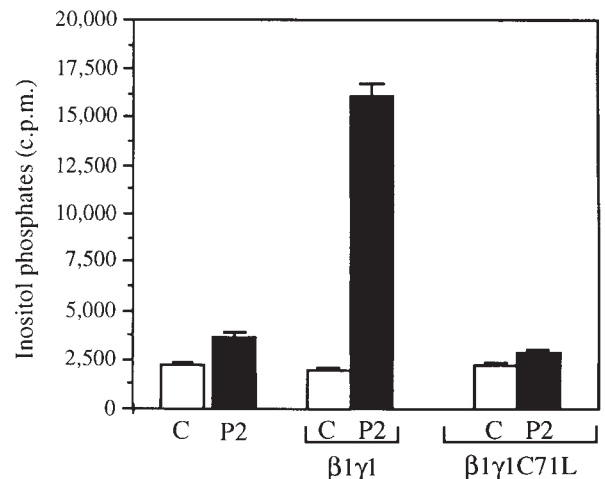


FIG. 2 Effect of $\beta_1\gamma_1$ C71L on PLC- β 2 in transfected COS-7 cells. Accumulation of inositol phosphates in COS-7 cells transfected with the indicated cDNAs; details and methods as for Fig. 1. Data represent duplicate determinations in a single experiment, and error bar shows the range of determinations. Three additional experiments gave similar results. The relative expression of each component in the various transfections was similar by western blotting.

led to increased inositol phosphate release, indicating that PLC- $\beta 1$ protein was synthesized and could be activated.

Activation is restricted to cells in which β - and γ -subunits are produced together. Thus, the G protein β_1 subunit or either of the γ_1 or γ_2 subunits when co-transfected on their own with PLC- $\beta 2$ had little or no effect on the accumulation of inositol phosphates (Fig. 1c). We conclude therefore that a complex of free $\beta\gamma$ subunits acts together to stimulate the activity of PLC- $\beta 2$ specifically. If free $G_{\beta\gamma}$ subunits are required for activity we should be able to neutralize this activity by providing sufficient $G\alpha$ to form an inactive heterotrimer. Co-transfection of cells with DNAs encoding $G_{\beta_1\gamma_1}$, phospholipase C- $\beta 2$ and $G_{\alpha_{12}}$ completely prevented activation of the phospholipase (Fig. 1c). Therefore excess $G_{\alpha_{12}}$ can act as a 'scavenger' of free $\beta\gamma$ complex and neutralize its activity.

These results indicate that specific activation of PLC- $\beta 2$ requires a free complex of the G protein $\beta\gamma$ subunits. The γ -subunit is modified by addition of an isoprenoid moiety (farnesyl or geranylgeranyl) at a cysteine residue four amino acids from the carboxy terminus of the nascent protein (the 'CAAX box'). In the absence of this modification, the $\beta\gamma$ complex forms but remains in the cytosol¹⁸⁻²⁰ and does not associate with the membrane. We have confirmed these observations (Fig. 1a) and have used mutants to test whether the $\beta\gamma$ complex that activates PLC has to be polyisoprenylated and membrane-associated. Figure 2 shows that although $\beta_1\gamma_1$ activates PLC- $\beta 2$, the mutant lacking the modification has no activity.

The activation of the PLC- $\beta 2$ isoform by free $\beta\gamma$ subunits may explain the pertussis-toxin-sensitive activation of PLC in some systems (reviewed in refs 13, 14). Pertussis toxin (PTx) ADP-ribosylates the α -subunit of certain heterotrimeric G proteins and uncouples the G protein from its receptor (reviewed in ref. 21). ADP-ribosylation may prevent receptor activation and subsequent release of $\beta\gamma$ subunits. Thus PTx could block activities mediated by free $\beta\gamma$. To test this idea, we used the M2 muscarinic acetylcholine (ACh) receptor because it inhibits adenylyl cyclase and opens a G_i -regulated K^+ channel by interacting with G_{α_i} (refs 22, 23); it also interacts with G_{α_i} in reconstitution experiments^{24,25}. In our system, transfection with the muscarinic ACh receptor alone shows little carbachol (ligand)-mediated release of inositol phosphates (Fig. 3). Co-transfection with PLC- $\beta 2$ generates a significant ligand-stimulated response which can be inhibited by pertussis toxin. This response presumably involves coupling of the transfected receptor to the recombinant PLC- $\beta 2$ by an endogenous G protein which is sensitive to PTx. In control experiments with the M1

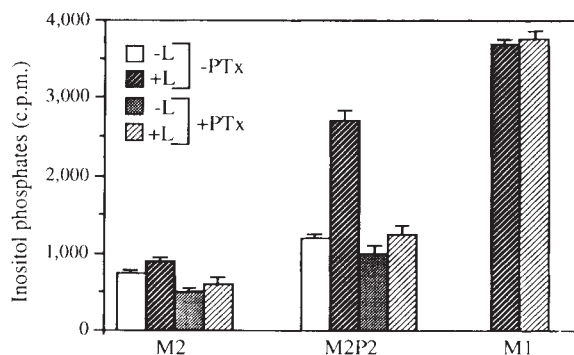


FIG. 3 Carbachol-induced accumulation of inositol phosphates in COS-7 cells transfected with cDNAs indicated with or without pertussis toxin (PTx). Cells were treated with PTx (400 ng ml^{-1}) 4 h before addition of $10 \mu\text{M}$ carbachol (L) during the last 25 min of the assay. M1 and M2 correspond to cDNAs of muscarinic acetylcholine receptor subtypes 1 and 2 respectively, both in pCIS. Details and methods as for Fig. 1. Cells were transfected with $0.5 \mu\text{g}$ of each cDNA. Data represent duplicate determinations in a single experiment, and error bar gives the range of the determinations; three additional experiments gave similar results.

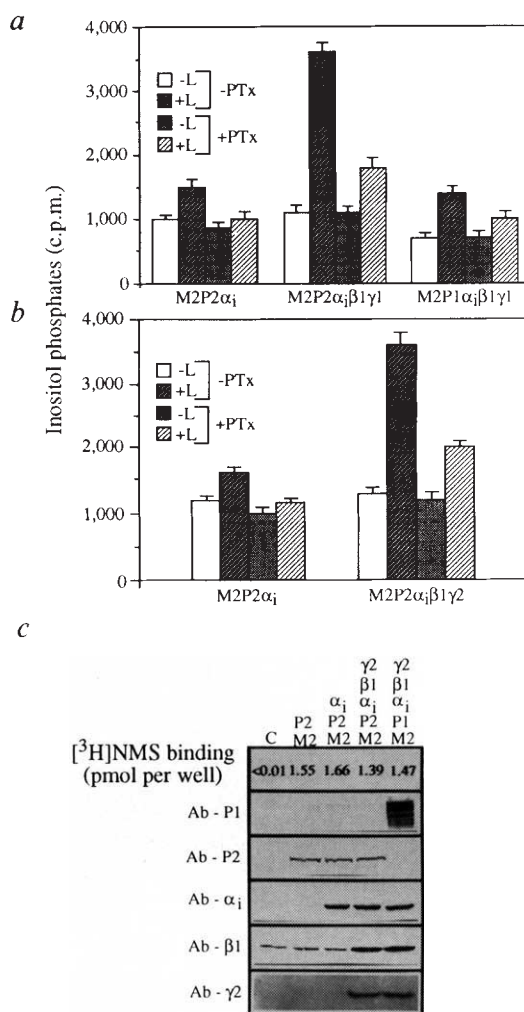


FIG. 4 Carbachol-induced accumulation of inositol phosphates in COS-7 cells transfected with the indicated cDNAs with or without pertussis toxin (PTx) treatment. Cells were treated with PTx (400 ng ml^{-1}) for 4 h before addition of $10 \mu\text{M}$ carbachol (L) which was added during the last 25 min of the assay. **a** and **b**, Cells were transfected with $0.2 \mu\text{g}$ each cDNA. Details and methods as in legends to Figs 1 and 3. Data represent duplicate determinations in a single experiment; three additional experiments gave similar results. **c**, Analysis of the relative expression levels of the transfected cDNAs corresponding to: M2, P1, P2, G_{α_i} (α_i), G_{β_1} and G_{γ_2} in the various multicomponent transfections. Transfected cDNAs are indicated above the panel. On the left are indicated the binding of the muscarinic antagonist [^3H]-methyl scopolamine ([^3H]NMS) and the antibodies (Ab) used for monitoring the levels of each component. The relative expression of each component was similar in the indicated transfections. M2 receptor expression was determined by saturation binding with [^3H]NMS. The relative expression of P1, P2, α_i , β_1 and γ_2 were determined by probing western blots with specific antibodies (P1 is partially degraded, its intact form is represented by the highest band). **METHODS.** COS-7 cells (1×10^5 cells per well) were transfected with $0.2 \mu\text{g}$ of each cDNA and pCIS encoding β -galactosidase was used to maintain the total DNA at $1 \mu\text{g}$. Saturation binding of [^3H]NMS to transfected COS-7 cells was done 48 h post-transfection with 5 concentrations of [^3H]NMS (0.1–23 nM). Cells were rinsed twice with PBS/0.1% BSA and incubated with 0.5 ml PBS/0.1% BSA containing labelled ligand for 1 h on ice. Cells were then rinsed four times with cold PBS/0.1% BSA and solubilized with 0.5 ml of 1 M NaOH. Bound [^3H]NMS was quantified by liquid scintillation counting. Scatchard analysis of saturation binding gave [^3H]NMS binding in pmol per well. Western blots were done as described in Fig. 1 legend but with unfractionated cell extracts. Equal amounts of extracts were resolved by SDS-PAGE (7.5% acrylamide for PLC- $\beta 1$ and PLC- $\beta 2$; 12% for G_{α_i} and G_{β_1} ; and 15% for G_{γ_2}). Antibodies against PLC- $\beta 1$ and PLC- $\beta 2$ were against the entire protein and do not crossreact. The G_{α_i} antibody was an antipeptide antibody directed against the C-terminal 10 amino acids. G_{β_1} and G_{γ_2} antibodies are described in Fig. 1 legend.

muscarinic receptor subtype, which couples to endogenous PLC- $\beta 1$ through the PTx-insensitive G_{α_q} and $G_{\alpha_{11}}$ proteins (ref. 26; A.K. and M.I.S., unpublished results), the response was unaffected by PTx (Fig. 3). Addition of 10 μ M atropine completely blocked all carbachol-induced effects (data not shown).

To test for complete reconstitution of the system, cells were co-transfected with cDNAs corresponding to receptor, phospholipase and the heterotrimeric G protein components $G_{\alpha_{12}}$, G_{β_1} and G_{γ_1} or G_{γ_2} . Cells transfected with PLC- $\beta 2$ and all of the other components show a strong response to carbachol that is inhibited by PTx (Fig. 4a). Replacing G_{γ_1} with G_{γ_2} gave similar results (Fig. 4b); $G_{\alpha_{13}}$ also gave results similar to $G_{\alpha_{12}}$. On the other hand, there is little additional ligand-dependent activation when PLC- $\beta 1$ cDNA is substituted with that for PLC- $\beta 2$ and there is very little, if any, effect from PTx treatment (Fig. 4a). A control experiment (Fig. 4c) demonstrates that the relative expression levels of each transfected cDNA is not effected by multiple cotransfection. These results are all consistent with the hypothesis that ligand binding to the M2 muscarinic receptor can stimulate a PTx-sensitive G protein like $G_{\alpha_{12}}$ and release $\beta\gamma$ subunits that activate PLC- $\beta 2$. PTx treatment uncouples the G protein from the receptor and presumably prevents ligand-mediated release of $\beta\gamma$ subunits.

Our co-transfection assay results show that free $G_{\beta\gamma}$ subunits specifically activate the $\beta 2$ isoform of PLC and not the $\beta 1$ isoform. Unlike the modulating effects of $\beta\gamma$ on different isoforms of adenylyl cyclase, there is no apparent requirement for an activated G_{α} subunit. A G_{α} protein that we find effectively stimulates PLC- $\beta 2$ is $G_{\alpha_{16}}$, however it is not found in Cos-7 cells and is PTx-insensitive^{6,7}. The observation that $G_{\beta\gamma}$ activates PLC- $\beta 2$, together with the fact that PLC- $\beta 2$ was cloned from HL60 cells⁷, suggests that the PTx-sensitive PLC-activation pathways in HL-60 cells^{13,14} could involve $\beta\gamma$ release. These results are consistent with *in vitro* reconstitution experiments¹² in which phosphatidylinositol PLC activity from HL-60 cell extracts was found to be activated by $G_{\beta\gamma}$; this probably corresponds to the $\beta 2$ isoform.

These results suggest that ligand binding to one particular receptor can stimulate one effector pathway through the G-protein-associated α -subunit and a different pathway through $G_{\beta\gamma}$. Complex circuits of intracellular responses can thus be generated as a result of receptor occupancy (reviewed in ref. 27). The different isoforms of the PLC- β subfamily are tissue-specific in their expression, so in some differentiated cell types the same heterotrimeric G proteins can activate different effectors. These observations extend our view of the versatility of the G-protein system and point to its role in coordinating complex responses to hormones and other active ligands in differentiated cells. □

Received 12 June; accepted 27 October 1992.

- Gilman, A. G. *Rev. Biochem.* **56**, 615-649 (1987).
- Birnbaumer, L., Abramowitz, J. & Brown, A. M. *Biochem. biophys. Acta* **1031**, 163-224 (1990).
- Smrcka, A. V., Hepler, J. R., Brown, K. O. & Sternweis, P. C. *Science* **251**, 804-807 (1991).
- Taylor, S. J., Chae, H. Z., Rhee, S. G. & Exton, J. H. *Nature* **350**, 516-518 (1991).
- Wu, D., Lee, C. H., Rhee, S. G. & Simon, M. I. *J. Biol. Chem.* **267**, 1811-1817 (1992).
- Lee, C. H., Park, D., Wu, D., Rhee, S. G. & Simon, M. I. *J. Biol. Chem.* **267**, 16044-16047 (1992).
- Park, D., Jhon, D.-Y., Kriz, R., Knopf, J. & Rhee, S. G. *J. Biol. Chem.* **267**, 16048-16055 (1992).
- Logothetis, D. E., Kurachi, Y., Galper, J., Neer, E. J. & Clapham, D. E. *Nature* **325**, 321-326 (1987).
- Jetsma, C. L. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3623-3627 (1987).
- Haga, K. & Haga, T. *J. Biol. Chem.* **267**, 2222-2227 (1992).
- Tang, W.-J. & Gilman, A. G. *Science* **254**, 1500-1503 (1991).
- Camps, M. et al. *Eur. J. Biochem.* **206**, 821-831 (1992).
- De Vivo, M. & Gershengorn, M. C. in *ADP-Ribosylating Toxins and G Proteins* (eds Moss, J. & Vaughan, M.) 267-293 (Am. Soc. Microbiol., Washington DC, 1990).
- Snyderman, R., Perlanin, A., Evans, T., Polakis, P. & Didsburg, J. in *ADP-Ribosylating Toxins and G Proteins* (eds Moss, J. & Vaughan, M.) 295-323 (Am. Soc. Microbiol., Washington DC, 1990).
- Federman, A. D., Conklin, B. R., Schrader, K. A., Reed, R. R. & Bourne, H. R. *Nature* **356**, 159-161 (1992).
- Cotecchia, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7159-7163 (1988).
- Osawa, S., Dhanasekaran, N., Woon, C. W. & Johnson, G. L. *Cell* **63**, 697-706 (1990).
- Prinin, A. N. & Gautam, N. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6220-6224 (1992).
- Simonds, W. F., Butrynyski, J. E., Gautam, N., Unson, C. G. & Spiegel, A. M. *J. Biol. Chem.* **266**, 5363-5366 (1991).
- Muntz, K. H., Sternweis, P. C., Gilman, A. G. & Mumby, S. M. *Molec. Biol. Cell* **3**, 49-61 (1992).
- Uj, M. in *ADP-Ribosylating Toxins and G Proteins* (eds Moss, J. & Vaughan, M.) 45-77 (Am. Soc. Microbiol., Washington DC, 1990).
- Ashtkenazi, A. et al. *Science* **238**, 672-675 (1987).

- Yatani, A., Codina, J., Brown, A. M. & Birnbaumer, L. *Science* **235**, 207-211 (1987).
- Tota, M. R., Kahler, K. R. & Schimerlik, M. I. *Biochemistry* **26**, 8175-8182 (1987).
- Parker, E. M., Kameyama, K., Higashijima, T. & Ross, E. M. *J. Biol. Chem.* **266**, 519-527 (1991).
- Berstein, G. et al. *J. Biol. Chem.* **267**, 8081-8088 (1992).
- Simon, M., Strathmann, M. & Gautam, N. *Science* **252**, 802-808 (1991).
- Bonifacio, J. S., Suzuki, C. K. & Klausner, R. D. *Science* **247**, 79-82 (1990).
- Gorman, C. M., Gies, D. R. & McCray, G. *DNA Protein Engn Tech.* **2**, 3-10 (1990).
- Berridge, M. J., Dawson, R. M. C., Downes, P., Heslop, J. P. & Irvine, R. F. *Biochem. J.* **212**, 473-482 (1983).

ACKNOWLEDGEMENTS. A.K. and D.W. contributed equally to this work. We thank S. G. Rhee, J. Knopf, E. Peralta, W. Simonds and R. Perlmutter for cDNA clones and antisera. This work is supported by a US public Health Services grant to M.I.S.

Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases

M. Rozakis-Adcock*, J. McGlade*, G. Mbamalu*, G. Pelicci†, R. Daly‡, W. Li‡, A. Batzer‡, S. Thomas§, J. Brugge§, P. G. Pelicci†, J. Schlessinger‡, & T. Pawson*||

* Division of Molecular and Developmental Biology, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada

† Istituto Clinico Medica 1, Policlinico Monteluce, University of Perugia, 06100 Perugia, Italy

‡ Department of Pharmacology, New York University Medical Center, 550 First Avenue, New York 10016, USA

§ ARIAD Pharmaceuticals Inc., 26 Landsdowne Street, Cambridge, Massachusetts 02139, USA

|| Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

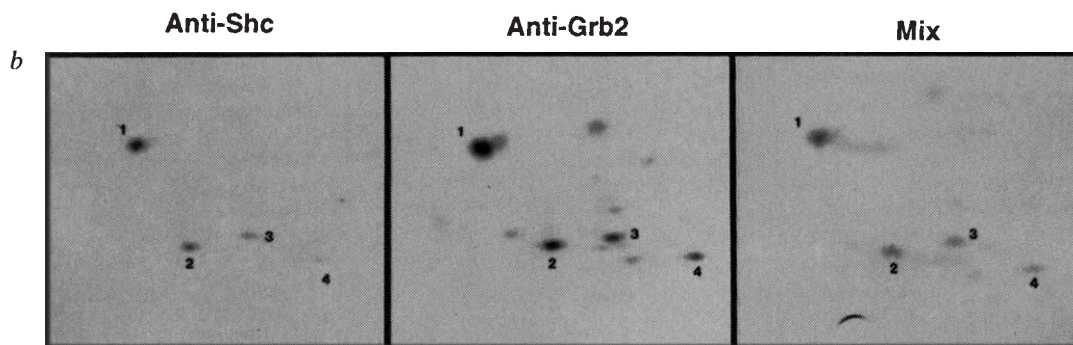
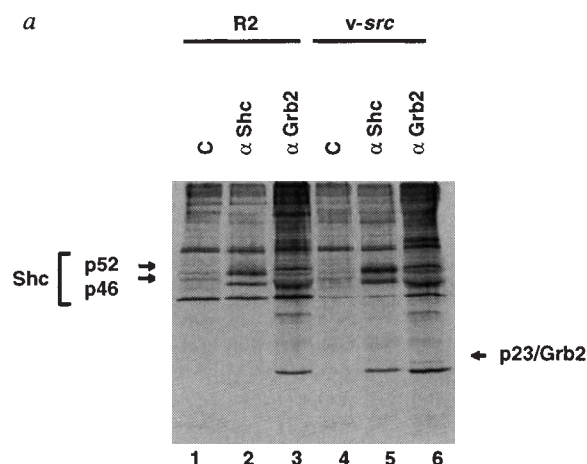
THE mammalian *shc* gene encodes two overlapping, widely expressed proteins of 46 and 52K, with a carboxy-terminal SH2 domain that binds activated growth factor receptors, and a more amino-terminal glycine/proline-rich region¹. These *shc* gene products (Shc) are transforming when overexpressed in fibroblasts¹. Shc proteins become phosphorylated on tyrosine in cells stimulated with a variety of growth factors¹, and in cells transformed by *v-src* (ref. 2), suggesting that they are tyrosine kinase targets that control a mitogenic signalling pathway. Here we report that tyrosine-phosphorylated Shc proteins form a specific complex with a non-phosphorylated 23K polypeptide encoded by the *grb2/sem-5* gene^{3,4}. The *grb2/sem-5* gene product itself contains an SH2 domain, which mediates binding to Shc, and is implicated in activation of the Ras guanine nucleotide-binding protein by tyrosine kinases in both *Caenorhabditis elegans* and mammalian cells^{3,4}. Consistent with a role in signalling through Ras, *shc* overexpression induced Ras-dependent neurite outgrowth in PC12 cells. These results suggest that Shc tyrosine phosphorylation can couple tyrosine kinases to Grb2/Sem-5, through formation of a Shc-Grb2/Sem-5 complex, and thereby regulate the mammalian Ras signalling pathway.

In Rat-2 cells transformed by *v-src* or *v-fps*, and in epidermal growth factor (EGF)-stimulated cells, the 46 and 52K Shc proteins (p46^{shc}, p52^{shc}) become phosphorylated on tyrosine^{1,2}. A minor and variably expressed 66K protein, encoded by a distinct *shc* transcript¹, is also tyrosine phosphorylated. Metabolic labelling of these cells with ³⁵S-methionine and ³²P has shown that activated tyrosine kinases induce the association of Shc proteins with a 23K polypeptide (p23), which is not itself phosphorylated on tyrosine, but which coprecipitates with anti-Shc

¶ To whom correspondence should be addressed.

FIG. 1 The p23 proteins immunoprecipitated with anti-Shc and anti-Grb2 antibodies are structurally indistinguishable. **a**, Comparison of anti-Shc and anti-Grb2 immunoprecipitates from ^{35}S -methionine-labelled Rat-2 or Rat-2 *v-src* cells. Rat-2 cells (R2; lanes 1–3) and Rat-2 *v-src* cells (*v-src*; lanes 4–6) were metabolically labelled with ^{35}S -methionine lysed and immunoprecipitated with affinity-purified anti-Shc (αShc) antibodies (lanes 2, 5), anti-Grb2 (αGrb2) antiserum (lanes 3, 6) or rabbit anti-mouse immunoglobulin antibodies as a control (C) (lanes 1, 4). A fraction of the Shc proteins from Rat-2 *v-src* cells (lane 5) run with a reduced mobility owing to their tyrosine phosphorylation². **b**, Tryptic peptide maps of p23 immunoprecipitated from ^{35}S -methionine-labelled Rat-2 *v-src* cells with anti-Shc or anti-Grb2 antibodies. The ^{35}S -methionine-labelled p23 bands immunoprecipitated with anti-Shc antibodies (lane 5, **a**) or with anti-Grb2 antiserum (lane 6, **a**) were excised and digested with trypsin. Tryptic peptides from each sample were resolved either alone (Anti-Shc; Anti-Grb2) or together (MIX), in two dimensions by thin-layer electrophoresis and ascending chromatography. The four major spots (1–4) are indicated in each map; the minor spots were also in common.

METHODS. Cells were labelled with $150\ \mu\text{Ci ml}^{-1}\ ^{35}\text{S}$ -methionine for 16 h in methionine-free Dulbecco's modified Eagle medium (DMEM) containing 0.5% dialysed calf serum. Immunoprecipitations and analysis on 12% SDS-PAGE were done as outlined previously¹⁷. The Shc proteins in lane 6 are obscured by the immunoglobulin heavy chain from the anti-Grb2 antiserum. Tryptic peptide map analysis on thin-layer cellulose plates used electrophoresis in the first dimension at pH 2.1, followed by chromatography in the second dimension, as previously described^{18,19}. Plates were exposed to X-ray film for 3–4 weeks.



antibodies² (Fig. 1a, lanes 2 and 5). The *grb2* gene, a mammalian homologue of the *C. elegans sem-5* gene, encodes a 23–25K protein with the structure SH3–SH2–SH3, which is not phosphorylated on tyrosine but which coprecipitates with a phosphotyrosine-containing protein of ~55K from EGF-stimulated cells⁴. To examine whether the p23 polypeptide that inducibly associates with phosphorylated Shc proteins might be related to Grb2, parental Rat-2 fibroblasts, or *v-src*-transformed Rat-2 cells (Rat-2 *v-src*), were metabolically labelled with ^{35}S -methionine, and immunoprecipitated with anti-Shc or anti-Grb2 antibodies (Fig. 1a). The radiolabelled, Shc-associated p23 was compared by tryptic peptide analysis with the comigrating Grb2 protein, precipitated directly with anti-Grb2 antiserum. Two-dimensional tryptic peptide maps of the two polypeptides were indistinguishable (Fig. 1b). These results indicate that the p23 protein complexed with Shc proteins in *v-src*-transformed cells is the *grb2* gene product.

The possibility that Shc proteins physically interact with Grb2 was pursued by immunoprecipitating lysates of Rat-2 and Rat-2 *v-src* cells with either anti-Shc or anti-Grb2, followed by immunoblotting with antibodies to phosphotyrosine (Fig. 2a). Strikingly, the spectrum of phosphotyrosine-containing proteins in anti-Grb2 and anti-Shc immunoprecipitates from *v-src*-transformed cells appeared identical. Although Grb2 was not itself phosphorylated on tyrosine in Rat-2 *v-src* cells, it was associated with two prominent phosphotyrosine-containing proteins that comigrated with p46^{shc} and p52^{shc}. Immunoblotting of the same immunoprecipitates with anti-Shc antibodies confirmed that the 46 and 52K proteins complexed with Grb2 in Rat-2 *v-src* cells were indeed *shc* gene products (Fig. 2b). In a similar experiment,

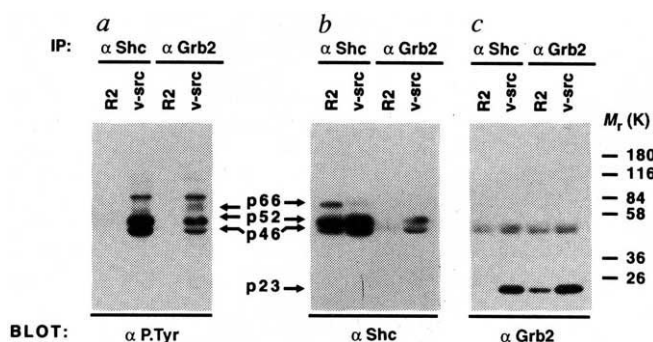
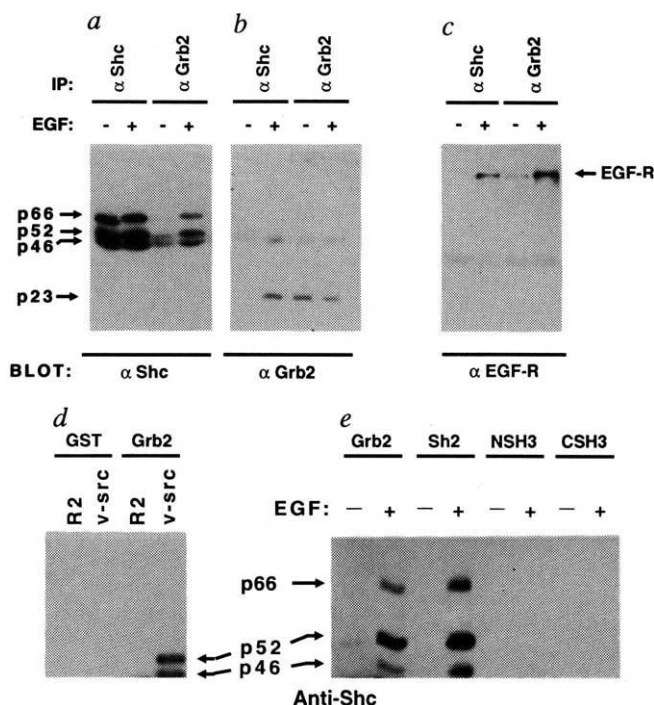


FIG. 2 Formation of a Shc–Grb2 complex in *v-src*-transformed Rat-2 cells. Cell lysates prepared from Rat-2 (R2) or Rat-2 *v-src* (*v-src*) cells were immunoprecipitated with anti-Shc and anti-Grb2 antibodies. The immunoprecipitates were subjected to immunoblot analysis with anti-phosphotyrosine ($\alpha\text{P.Tyr}$) antibodies (**a**), anti-Shc (αShc) antibodies (**b**), and anti-Grb2 (αGrb2) antibodies (**c**), followed by detection with ^{125}I -labelled protein A. The mobilities of the Shc proteins (p46, p52, p66) and of the Grb2 protein (p23) are indicated.

METHODS. Rat-2 and Rat-2 *v-src* cells, were lysed and immunoprecipitated as previously described^{2,20}. In each case 1 ml of cell-lysate ($\sim 0.8\ \text{mg ml}^{-1}$) was incubated for 90 min at 4°C with $1\ \mu\text{g}$ of affinity-purified anti-Shc polyclonal rabbit antibodies^{1,2}, or with $5\ \mu\text{l}$ of anti-Grb2 rabbit antiserum⁴. Washed immune complexes were separated on a 10% SDS-polyacrylamide gel and transferred onto nitrocellulose membranes. Blots were blocked and probed with anti-phosphotyrosine, anti-Shc, or anti-Grb2 antibodies as previously described^{2,4,20}. The background band at 50–55K is the immunoglobulin heavy chain from the immunoprecipitation.

FIG. 3 EGF-dependent association of Grb2 with the Shc proteins *in vivo* and *in vitro*. Rat-1 cells overexpressing the human EGF receptor (R1HER cells) were serum-starved (–), or stimulated with EGF(+). Cells were lysed and immunoprecipitated with anti-Shc or anti-Grb2 antibodies. Immunoprecipitates were subjected to Western blot analysis with anti-Shc (α Shc) antibodies (a), anti-Grb2 (α Grb2) antibodies (b), or anti-EGF-receptor (α EGF-R) antibodies (c). To analyse binding of Grb2 to tyrosine phosphorylated Shc proteins *in vitro*, bacterially expressed glutathione *S*-transferase (GST), and GST fusion protein containing the full-length human Grb2 polypeptide (Grb2), were immobilized with glutathione-agarose and incubated with lysates of Rat-1 cells (R2) or Rat-2 v-src cells (v-src), and analysed for associated Shc proteins by immunoblotting with anti-Shc antibodies (d). Alternatively, cells expressing the human EGFR were serum-starved (–) or stimulated with EGF(+), lysed and incubated with GST-fusion proteins corresponding to full-length Grb2 (Grb2), the SH2 domain of Grb2 (SH2), or the individual SH3 domains of Grb2 (denoted as NSH3 for N-terminal and CSH3 for C-terminal domains, respectively). Associated proteins were subsequently analysed by immunoblotting with anti-Shc antibodies (e).

METHODS. a–c, R1HER cells were stimulated for 5 min with 80 nM EGF, following 16 h of serum starvation in DMEM containing 0.5% calf serum, before immunoprecipitation and blotting. d, e, Serum-starved HER-14 cells²¹ were stimulated for 2 min with 275 ng ml^{–1} EGF. Clarified cell lysates were incubated with ~5 μ g of GST fusion proteins for 90 min at 4 °C. Preparation of bacterial GST-fusion proteins and *in vitro* binding to tyrosine phosphorylated proteins in cell lysates has been previously described^{1,4,20}. The washed complexes were separated by SDS-PAGE (8.25%), transferred to nitrocellulose and blotted with anti-Shc antibody. The mobilities of the Shc proteins (p46, p52, p66), Grb2 (p23) and the EGF-receptor (EGF-R) are indicated.



the 23K Shc-associated protein, specifically immunoprecipitated from Rat-2 v-src cells with anti-Shc antibodies, was directly recognized by antibodies to Grb2 in an immunoblot (Fig. 2c).

These results support the contention that tyrosine phosphorylation of Shc proteins by the v-Src tyrosine kinase induces the formation of a Shc–Grb2 complex. Similar results were obtained using Rat-2 cells transformed by the v-Fps tyrosine kinase (data not shown).

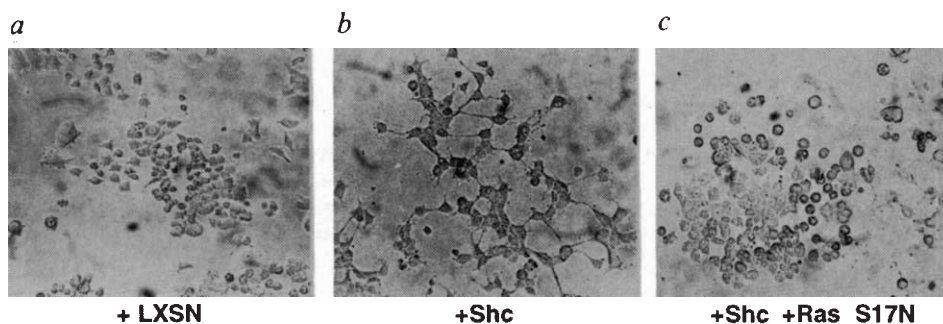
The association of Shc and Grb2 in cells transformed by oncogenic tyrosine kinases raised the possibility that this interaction might be a facet of the normal cellular response to mitogenic growth factors. We therefore analysed Shc and Grb2 in rodent fibroblasts expressing the human EGF-receptor (EGFR) (Fig. 3). In serum-starved cells, Shc proteins were only weakly bound to Grb2. However, EGF stimulation induced the formation of a Shc–Grb2 complex (Fig. 3a, b), concomitant with Shc tyrosine phosphorylation (data not shown). In addition to Shc and Grb2, the EGFR could be detected in anti-Shc or anti-Grb2 immunoprecipitates from EGF-stimulated cells^{1,2,4} (Fig. 3c). Activation of the EGFR therefore induces the formation of hetero-oligomeric complexes containing the autophosphorylated EGFR, and the Shc and Grb2 SH2-containing proteins. The tyrosine phosphorylation of Shc proteins, and their association with Grb2, is apparently a common feature of cells either

transformed by oncogenic tyrosine kinases, or stimulated with EGF.

To explore the basis for the interaction between Shc and Grb2 proteins, we attempted to recreate this complex *in vitro*, using the full-length Grb2 protein expressed in bacteria as a glutathione *S*-transferase (GST) fusion protein (Fig. 3d, e). The immobilized bacterial Grb2 did not bind Shc proteins in a lysate of normal Rat-2 cells, but stably associated with p46^{shc} and p52^{shc} in a lysate of v-src-transformed Rat-2 cells. Similarly, the GST–Grb2 fusion bound all three Shc proteins in a lysate of EGF-stimulated cells, but not in unstimulated cells. A GST fusion protein containing only the Grb2 SH2 domain was also effective in binding Shc proteins from EGF-stimulated cells, whereas bacterial fusion proteins containing either SH3 domain alone showed no *in vitro* Shc-binding activity (Fig. 3e). Consistent with these observations, GST–Grb2 fusion proteins with amino-acid substitutions in the SH2 domain, corresponding to loss-of-function *sem-5* mutations^{3,4}, were impaired in their association with Shc proteins, whereas SH3 substitutions were without effect (data not shown). It appears that the interaction between Shc and Grb2 is mediated by binding of the Grb2 SH2 domain to tyrosine phosphorylated Shc. NIH 3T3 cells overexpressing Shc proteins that lack the presumptive tyrosine phosphorylation site are not transformed (A. E. Salcini, G.P. and

FIG. 4 *shc* overexpression induces Ras-dependent neurite outgrowth in PC12 cells. PC12 cells, modified to express inducibly the dominant inhibitory RasS17N mutant under the control of the MMTV promoter¹⁰ (GSrasDN6 cells), were infected with the retrovirus vector LXS^N (a), or with a modified retrovirus (L-SHC-SN) encoding the P46^{shc} and p52^{shc} proteins (b, c). Cells in c were treated with 150 nM dexamethasone for 12 h before infection, and for a further 48 h to induce expression of the Ras S17N protein. Similar results were obtained when dexamethasone was included throughout the infection. Photomicrographs were taken 7 days after infection.

METHODS. The human *shc* cDNA was cloned into the *Eco*RI site of LXS^N²², and high titre helper-free virus was produced, essentially as described²³.



GSrasDN6 cells were infected in the presence of 6 μ g ml^{–1} polybrene. Infected cultures overexpressed Shc proteins, as determined by immunoblotting. Cells were cultured and induced with dexamethasone, essentially as described previously¹⁰.

P.G.P., unpublished results). This is consistent with the notion that association with Grb2 is important for Shc transforming activity.

In *v-src*-transformed cells, ^{35}S -methionine-labelling has not revealed stoichiometric amounts of a third protein in Shc-Grb2 complexes. Because Grb2 is not itself a tyrosine kinase substrate, tyrosine phosphorylation of Shc may be the principal mechanism by which non-receptor tyrosine kinases couple to Grb2, and hence to the Ras pathway. In EGF-stimulated cells, the EGFR can apparently bind stably to both SH2-containing proteins; this may precede Shc tyrosine phosphorylation and formation of any direct Shc-Grb2 complex.

Because the *grb2/sem-5* gene product lies upstream of Ras, these results predict that Shc proteins are also involved in control of the Ras signalling pathway³. To test this possibility, we have used the PC12 pheochromocytoma neuronal cell line. Stimulation of PC12 cells with nerve growth factor (NGF) activates the Trk tyrosine kinase⁵⁻⁷, and thereby induces neurite extension, in a response that is dependent on the Ras pathway⁸⁻¹¹. NGF stimulation of PC12 cells induced the tyrosine phosphorylation of Shc proteins, and the formation of a Shc-Grb2 complex (data not shown). Overexpression of the human *shc* complementary DNA in PC12 cells induced neurite extension; furthermore, this effect of *shc* was blocked by expression of a dominant inhibitory H-Ras mutant, RasS17N (ref. 12; Fig. 4). These results support the suggestion that Shc proteins are involved, through Grb2, in activation of the Ras pathway by tyrosine kinases. These findings do not exclude the possibility that other mechanisms exist for coupling tyrosine kinases to Grb2.

Although the biochemical function of the Shc-Grb2 complex is not known, a likely possibility involves interactions through their non-SH2 regions with proteins that directly regulate Ras, such as guanine nucleotide-release factors or GTPase-activating

proteins¹³⁻¹⁶. The formation of a complex involving the Shc and Grb2 SH2-containing proteins in growth factor-stimulated cells and in cells transformed by *v-Src*, which itself contains an SH2 domain, argues that a network of phosphorylation-dependent protein-protein interactions controls the activation of the Ras signalling pathway by tyrosine kinases. □

Received 12 August; accepted 4 November 1992.

1. Pelicci, G. *et al.* *Cell* **70**, 93-104 (1992).
2. McGlade, J., Cheng, A., Pelicci, G., Pelicci, P. G. & Pawson, T. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8869-8873 (1992).
3. Clark, S. G., Stern, M. J. & Horvitz, H. R. *Nature* **356**, 340-344 (1992).
4. Lowenstein, E. J. *et al.* *Cell* **70**, 431-442 (1992).
5. Hempstead, B. L., Martin-Zanca, D., Kaplan, D. R., Parada, L. F. & Chao, M. V. *Nature* **350**, 678-683 (1991).
6. Kaplan, D. R., Hempstead, B. L., Martin-Zanca, D., Chao, M. V. & Parada, L. F. *Science* **252**, 554-558 (1991).
7. Klein, R., Jing, S., Nanfuri, V., O'Rourke, E. & Barbacid, M. *Cell* **65**, 189-197 (1991).
8. Hagag, N., Haleboua, S. & Viola, M. *Nature* **319**, 680-682 (1986).
9. Szeberenyi, J., Cai, H. & Cooper, G. M. *Molec. cell. Biol.* **10**, 5324-5332 (1990).
10. Thomas, S. M., DeMarco, M., D'Arcangelo, G., Haleboua, S. & Brugge, J. S. *Cell* **68**, 1031-1040 (1992).
11. Wood, K. W., Sarnecki, C., Roberts, T. M. & Blenis, J. *Cell* **68**, 1041-1050 (1992).
12. Feig, L. A. & Cooper, G. M. *Molec. cell. Biol.* **8**, 3235-3243 (1988).
13. Bourne, H. R., Sanders, D. A. & McCormick, F. *Nature* **349**, 117-127 (1991).
14. Martegani, E. *et al.* *EMBO J* **11**, 2151-2157 (1992).
15. Simon, M. A., Bowtell, D. D. L., Dodson, G. S., Lavery, T. R. & Rubin, G. M. *Cell* **67**, 701-716 (1991).
16. Bonfini, L., Karlovich, C. A., Dasgupta, C. & Banerjee, U. *Science* **255**, 603-606 (1992).
17. Moran, M. F., Polakis, P., McCormick, F., Pawson, T. & Ellis, C. *Molec. cell. Biol.* **11**, 1804-1812 (1991).
18. Liu, X. & Pawson, T. *Molec. cell. Biol.* **11**, 2511-2516 (1991).
19. Luo, K., Hurley, T. R. & Sefton, B. M. *Oncogene* **5**, 921-923 (1990).
20. McGlade, C. J. *et al.* *Molec. cell. Biol.* **12**, 991-997 (1992).
21. Honegger, A. M. *et al.* *Cell* **51**, 199-209 (1987).
22. Miller, A. D. & Rosman, G. J. *Biotechniques* **7**, 980-990 (1989).
23. Miller, A. D., Tauber, D. R. & Buttimore, C. *Somatic Cell molec. Genet.* **12**, 175-183 (1986).

ACKNOWLEDGEMENTS. This work was supported by grants from the Medical Research Council of Canada, the National Cancer Institute of Canada (NCIC) with funds from the Canadian Cancer Society and Bristol-Myers Squibb to T.P., by a Howard Hughes International Scholar Award to T.P., and by grants from RPR, Sugan and HFSP to J.S. J.M. is a fellow of the NCIC. T.P. is a Terry Fox Cancer Research Scientist of the NCIC.

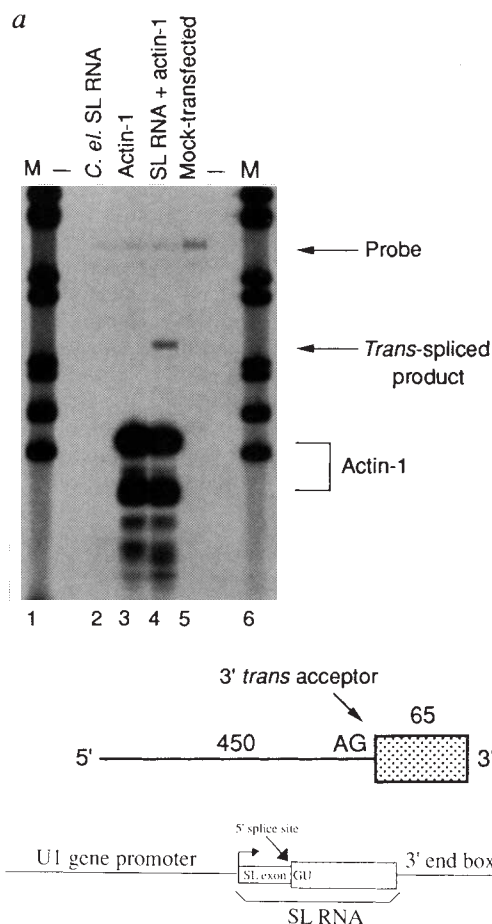
Spliced leader RNAs from lower eukaryotes are *trans*-spliced in mammalian cells

James P. Bruzik & Tom Maniatis

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

EXON sequences present on separate RNA molecules can be joined by *trans*-splicing in trypanosomatids, *Euglena*, and in the nematode and trematode worms¹⁻³. *Trans*-splicing involves an interaction between a 5' splice site present in a spliced leader RNA and a 3' splice site located near the 5' end of pre-messenger RNAs. *In vitro trans*-splicing of artificial mammalian pre-mRNAs has been reported, but the efficiency of splicing appears to depend on sequence complementarity between the two substrates⁴⁻⁷. There has been speculation that some natural pre-mRNAs can be *trans*-spliced in mammalian cells *in vivo*^{8,9}, but alternative interpretations have not been ruled out. Here we show that spliced leader RNAs can be accurately *trans*-spliced in mammalian cells *in vivo* and *in vitro*. Both nematode and mammalian 3' splice sites can function as acceptors for *trans*-splicing *in vivo*. These results reveal functional conservation in the splicing machinery between lower eukaryotes and mammals, and they directly demonstrate the potential for *trans*-splicing in mammalian cells.

*Caenorhabditis elegans*¹⁰ or *Leptomonas collosoma*¹¹ spliced leader (SL) RNAs were produced in mammalian cells using a well characterized human U1 small nuclear (sn) RNA expression vector (pUC-U1; ref. 12; Fig. 1a). These SL RNAs were accurately initiated and terminated, contained non-polyadenylated 3' termini, a trimethylguanosine cap, and were assembled into particles containing Sm proteins (data not shown). In order to



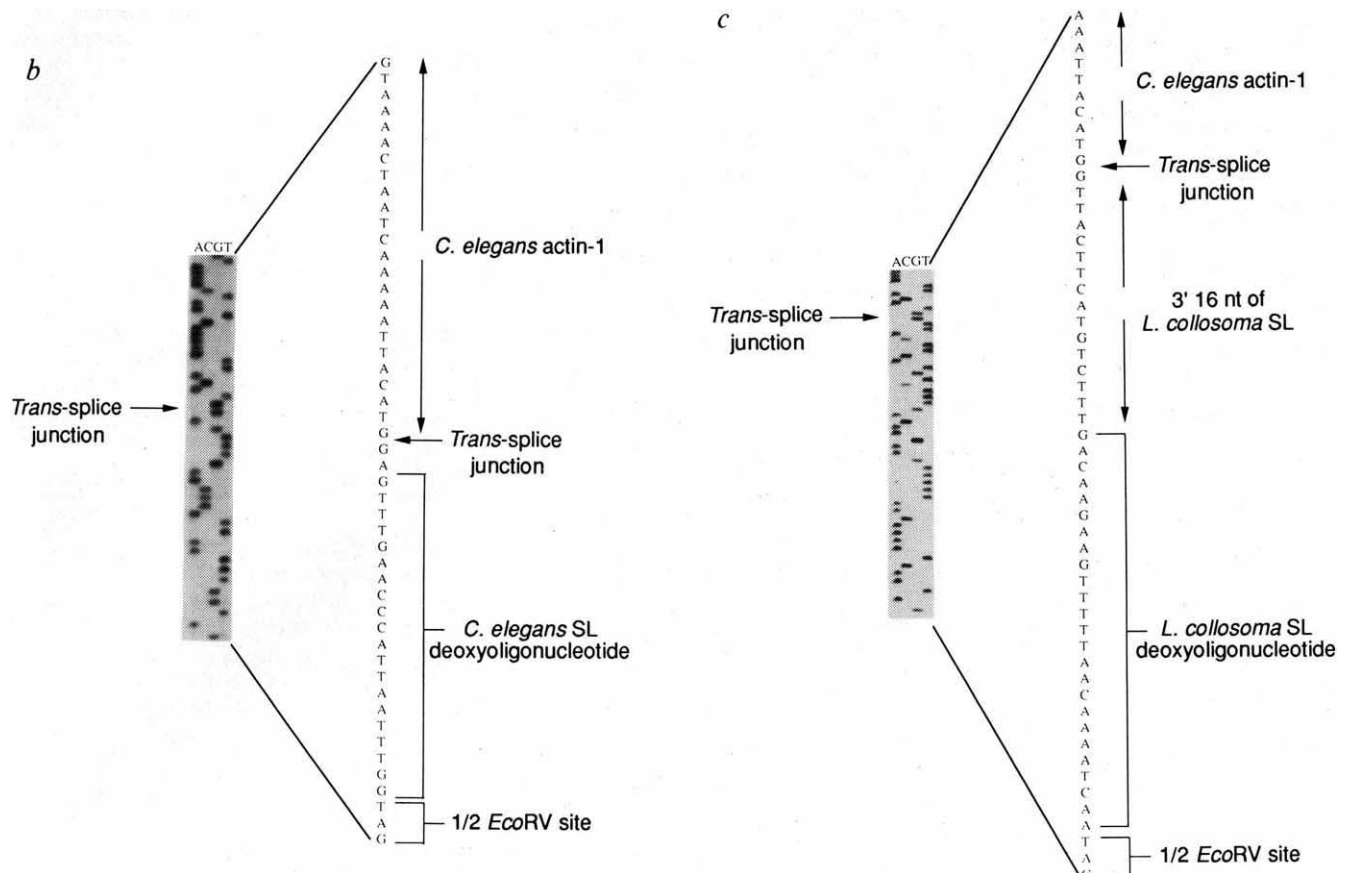


FIG. 1 Detection of *trans*-spliced RNAs *in vivo*. **a**, COS cells were transfected with either the *C. elegans* SL RNA or *C. elegans* actin-1 3' acceptor expression plasmids alone, or were cotransfected with both plasmids. RNA was isolated, selected with the *C. elegans* specific 2'-OCH₃ oligonucleotide, and analysed by RNase protection with a probe that spans the SL/actin-1 *trans*-splice junction. Lanes 2 and 3 show the protection of RNA from cells transfected with the SL RNA or actin-1 expression plasmids alone. When the two are cotransfected, a new band appears that corresponds to *trans*-spliced product (lane 4). Lane 5 shows the background of protection with this probe on RNA from mock-transfected COS cells. Lanes 1 and 6(M) contain pBR322 labelled *Msp* I-digested DNA size markers. The positions of the full-length probe, *trans*-spliced product, and actin-1 acceptor are denoted by arrows and a bracket. The stippled box represents actin-1 exon sequence and the 3' splice site is denoted by an arrow. SL exon is shown as an open box, the 3' region of the SL RNA as a hatched box and the 5' splice site is denoted by an arrow. **b**, Sequence of the *trans*-spliced product including the *C. elegans* SL RNA and actin-1 acceptor. From bottom to top, the sequence is seen to contain the *EcoRV* site, the 20-nucleotide *C. elegans* specific deoxyoligonucleotide, the next two bases of the SL exon up to the correct 5' splice site, and actin-1 sequence beginning at the correct 3' splice site. **c**, Sequence of the *trans*-spliced product using the *L. collosoma* SL RNA and nematode actin-1 acceptor *in vivo*. COS cells were transfected and the RNA was isolated and selected as before, except that the *L. collosoma* SL RNA gene and the *L. collosoma* SL-specific 2'-OCH₃ oligonucleotide were used. Here the sequence includes the *EcoRV* site, the 25-nucleotide *L. collosoma*-specific deoxyoligonucleotide, the next 16 bases of the SL exon up to the correct splice junction, followed by the actin-1 sequence starting at the 3' splice site.

METHODS. COS cells were grown in DMEM plus 10% fetal bovine serum and were transfected (50% confluent; 10-cm plate), with DEAE-dextran and collected after 48 h. Fragments of the *C. elegans* actin-1 gene were cloned into the vector pCDNA1 (Invitrogen) for *in vivo* expression. The actin-1 clone with more sequence 5' of the *trans* acceptor (**a**) was subcloned from pGem 3B (K. Van Doren and D. Hirsh) using an *EcoRI*-*Bam*HI fragment (515 bp) containing only actin-1 sequence, had the *Bam*HI site filled in and was ligated into *EcoRI*-*EcoRV*-digested pCDNA1. SL RNAs were expressed *in vivo* from the vector pUC U1, which contains a 681-bp *Hae*III fragment of a human U1 gene ligated to *Bam*HI linkers³¹, and cloned into the *Bam*HI site of pUC 1.3

(ref. 12). The U1 coding region was removed by digestion with *Bgl*II (position 2, relative to the 5' end of U1) and *Sal*I (downstream of the U1 insert). SL RNA clones were generated using two overlapping deoxyoligonucleotides as described¹⁵. Ends were engineered to contain *Bgl*II (5') and *Sal*I (3') restriction sites. The inserted DNA fragments precisely replace flanking bases that were removed with the U1 coding region. RNA was prepared as described³²; 2'-OCH₃ oligoribonucleotides (*C. elegans* SL exon and *L. collosoma* SL exon) were synthesized and biotinylated as described³³. Sequences are as follows: *C. el.* SL, 5'-CTCAAACCTGGGTAATTAAACBBBdG-3'; *Lept.* SL, 5'-GTTCTTCAAAAATTGTTTAGTTBBBdC-3'; B denotes biotinylated deoxycytidine residues. Binding reactions contained ~30 µg total RNA, 125 ng 2'-OCH₃ oligonucleotide (titrated for individual precipitations), 50 mM Tris, pH 8.3, 60 mM NaCl and 10 mM DTT. They were heated to 65 °C for 5 min, incubated at 37 °C for 30 min and placed on ice. Streptavidin-agarose beads (Sigma; 20 µl packed beads per reaction) were pre-blocked and incubated with the binding reaction for 90 min at 4 °C with rocking as described³⁴, except that incubation and wash buffer contained 300 mM NaCl. For RNase protection³² a probe was used that spans the *C. elegans* SL/actin-1 *trans*-splice junction. Protected fragments were separated on a 6% polyacrylamide-7M urea gel and visualized by autoradiography. RT-PCR was done using specific deoxyoligonucleotides. RNA and 3' primer were incubated together at 65 °C for 5 min and then cooled on ice before reverse transcription reactions (20 µl) using avian myeloblastosis virus (AMV) reverse transcriptase (Promega; protocol from the supplier). PCR (100 µl) containing 1 µl reverse transcription reaction was done with AmpliTaq DNA polymerase (Cetus; protocol from the supplier). Reactions were cycled 35 times at 94 °C for 1 min, at 40 °C for 2 min, at 72 °C for 2 min, followed by extension at 72 °C for 6 min. Reaction products were purified and cloned into the *EcoRV* site of Bluescript KS⁺ (Stratagene). Inserts of the correct size were then dideoxynucleotide-sequenced in both directions with the T7 and T3 promoter primers and run on 6% polyacrylamide-7M urea gels and visualized by autoradiography. Specific primers for RT-PCR are as follows: *C. elegans* SL 5' primer, 5'-GGTTTAATTACCAAGTTTG-3' (20 nt); *L. collosoma* SL 5' primer, 5'-AACTAAACAATTTTGAAGAAGACAG-3' (25 nt); *in vivo* acceptor 3' primer (SP6 promoter primer, all *in vivo*-transcribed 3' acceptor fragments contained the SP6 promoter sequence at their 3' ends), 5'-CTAAATCCACTGTGATATC-3' (19 nt).

detect *trans*-spliced RNA we used a 2'-OCH₃ oligoribonucleotide-affinity selection procedure^{13,14} to enrich for molecules containing spliced leader exon sequences. Both full-length SL RNA and any product of the *trans*-splicing reaction containing this sequence should be selected. The specificity of the selection was demonstrated by showing that virtually all of the *C. elegans* SL RNA produced in COS cells could be selected with the 2'-OCH₃ oligonucleotide, whereas none was selected in the absence of the oligonucleotide (data not shown).

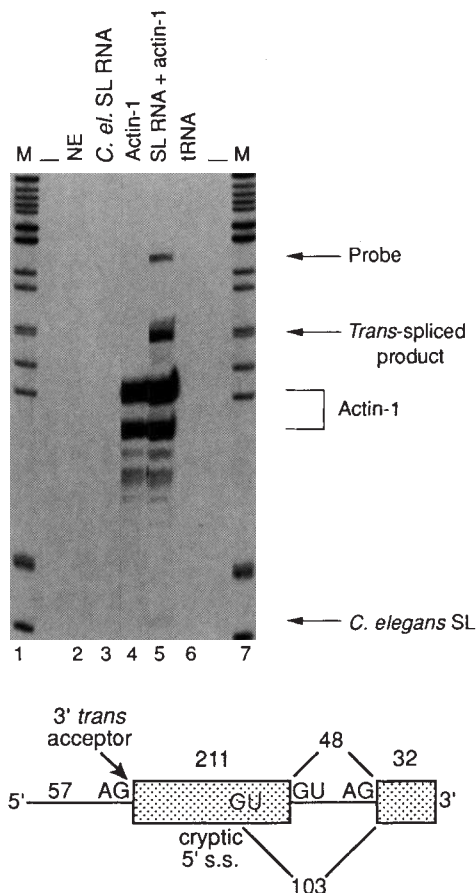


FIG. 2 *Trans*-splicing *in vitro* in HeLa cell nuclear extract. *C. elegans* SL RNA and actin-1 acceptors were transcribed *in vitro*, gel-purified, and incubated together in HeLa cell nuclear extract. Reactions were terminated after 2 h and the RNA was selected and analysed by RNase protection with a probe that spans the *C. elegans* SL/actin-1 junction. The background with HeLa nuclear extract and *E. coli* tRNA are shown in lanes 2 and 6. Addition of the SL RNA to the extract generates protection of a short fragment corresponding to the protection of the SL exon (lane 3). Lane 4 illustrates the protection obtained on incubation of actin-1 RNA (fragment containing more sequence 3' of the *trans*-splicing acceptor, although both fragments (Fig. 1a) serve as acceptors *in vitro* in nuclear extract. When both RNAs are incubated together, a new protected fragment appears, corresponding to *trans*-spliced product (lane 5). Lanes 1 and 7 contain pBR322 *Msp*I-labelled DNA size markers (M). Full-length probe, *trans*-spliced product and SL exon fragments are denoted by arrows, and actin-1 by a bracket. Stippled boxes represent actin-1 exons, lines represent intron sequence and the 3' *trans* acceptor as well as the cryptic internal 5' splice site are noted.

METHODS. *In vitro trans*-splicing used the *Hph*I fragment of actin-1 (clone with more 3' sequence) cloned into the filled-in *Bam*HI site of pGem 3B (K. Van Doren and D. Hirsh), and SL RNA clones described previously³⁵. *In vitro* reactions (25 μ l) were as described³⁶, except that they contained HeLa cell nuclear extract. SL RNA transcripts were incubated alone for 30 min at 30 °C after which 3' acceptor fragments were added and incubated for 2 h at 30 °C. Reactions were stopped by addition of 300 μ l NET-2 (50 mM Tris, pH 7.5, 150 mM NaCl, 0.05% NP40) plus 1% SDS, extracted with phenol/chloroform/isoamylalcohol (50:49:1) and precipitated with ethanol. Selection and RNase protection are described in Fig. 1 legend. The protected fragments were separated on an 8% polyacrylamide-7M urea gel and are visualized by autoradiography.

To determine whether the nematode SL RNA could be *trans*-spliced to an authentic nematode 3' *trans* acceptor in mammalian cells, COS cells were cotransfected with the SL RNA expression plasmid and the mammalian expression vector pcDNA 1 (Invitrogen) containing a fragment of the *C. elegans* actin-1 gene¹⁰ (Fig. 1a). When the SL RNA gene was transfected alone, no *trans*-spliced RNA was detected by RNase protection (lane 2; a short RNA fragment corresponding to the spliced leader exon was run off from this gel). Similarly, no *trans*-spliced RNA was detected in cells transfected with the actin-1 gene plasmid alone (lane 3). But when both the SL RNA and actin-1 RNA were expressed in COS cells, *trans*-spliced product was observed (lane 4). The sequence of the *trans*-spliced product was determined by DNA sequence analysis of double-stranded complementary DNA generated using reverse transcription followed by the polymerase chain reaction (RT-PCR) (see bottom, Fig. 1b). Parallel experiments were carried out with the *L. collosoma* trypanosomatid SL RNA and the *C. elegans* actin-1 *trans* acceptor. DNA sequence analysis of the *trans*-spliced product by RT-PCR revealed that the trypanosomatid SL RNA is also accurately *trans*-spliced in mammalian cells (Fig. 1c). *Trans*-splicing *in vivo* was also observed with both SL RNAs and an actin-1 acceptor containing an internal intron (see bottom of Fig. 2). Interestingly, this intron is removed through the use of the authentic 3' splice site and a cryptic 5' splice site 55 nucleotides upstream of the authentic 5' splice site (data not shown). Thus the same RNA molecule can be both *cis*- and *trans*-spliced in mammalian cells as is the case with endogenous *C. elegans* messages¹⁰.

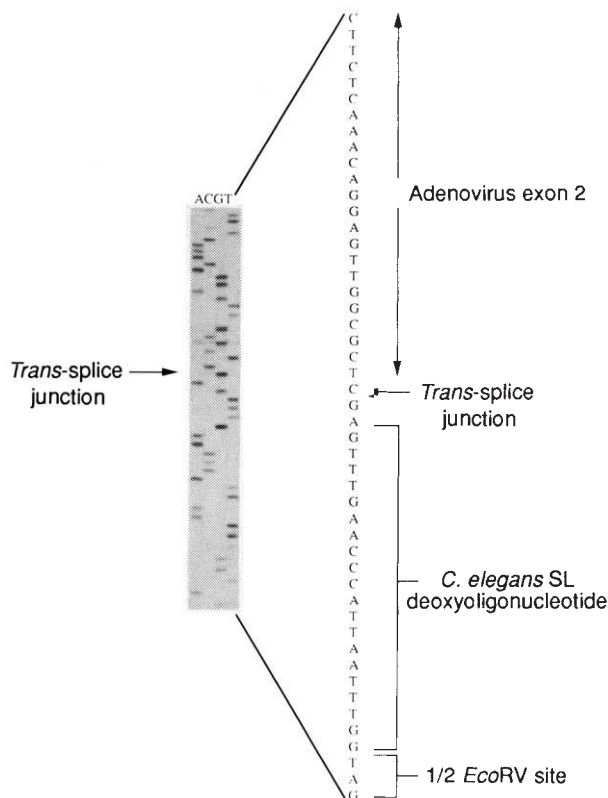


FIG. 3 Sequence of RNA generated by *trans*-splicing of the *C. elegans* SL RNA and the adenovirus major late intron-1 3' splice site *in vivo*. The *Eco*RV site is followed by the 20-nucleotide *C. elegans* specific primer, the next two bases of the SL exon, and crosses into the adenovirus exon-2 at the correct splice junction.

METHODS. All procedures were as described in Fig. 1 legend, except that the adenovirus major late transcript intron-1 half fragment containing the 3' splice site was subcloned as described¹⁵, generating a *Hind*III-*Hae*III fragment that was ligated into *Hind*III-*Eco*RV-digested pcDNA 1 for expression in mammalian cells.

We also carried out experiments to determine whether RNAs synthesized from separate plasmids could be *trans*-spliced *in vitro*. SL RNA and actin-1 pre-mRNAs were separately synthesized *in vitro*, gel-purified and then added either alone or in combination to a HeLa cell nuclear extract. The products were then selected with the *C. elegans* spliced-leader-specific 2'-OCH₃ oligonucleotide and assayed by RNase protection (Fig. 2). When SL RNA alone was added to the extract, a short protected fragment corresponding to spliced leader exon was observed (lane 3). Similarly, a protected RNA fragment corresponding in size to unspliced RNA was observed when *in vitro*-synthesized actin pre-mRNA was added to the extract (lane 4). But, like the situation *in vivo*, a protected band corresponding to the size of *trans*-spliced product was observed when both the SL RNA and 3' *trans*-splicing acceptor were added to the *in vitro* reaction (lane 5). The DNA sequence of the RT-PCR products with both *C. elegans* and *L. collosoma* SL RNAs and the actin-1 acceptor from *in vitro* reactions confirmed the generation of accurately *trans*-spliced products (data not shown). These results argue against the possibility that the *trans*-splicing observed *in vivo* was due to a low level of recombination between the cotransfected plasmids followed by *cis* splicing. They also establish the feasibility of studying *trans*-splicing in mammalian extracts.

Previous studies demonstrated that the nematode SL RNA can be spliced to an adenovirus pre-mRNA 3' splice site *in vitro*, when the two splice sites are present in the same RNA (in *cis*)¹⁵. To determine whether this mammalian 3' splice site can also participate in *trans*-splicing, the SL RNA expression vector was cotransfected with a vector containing the adenovirus 3' splice site. As with the nematode 3' splice site, *trans*-spliced product was observed only when the two plasmids were cotransfected (data not shown). The DNA sequence of the RT-PCR product from 2'-OCH₃-selected RNA revealed that the SL RNA is accurately *trans*-spliced to the adenovirus 3' splice site (Fig. 3).

The demonstration of *trans*-splicing of nematode and trypanosomatid SL RNAs in mammalian cells provides an additional evolutionary link between *cis*- and *trans*-splicing¹⁶, and splicing in lower and higher eukaryotes. All of these splicing reactions, as well as the self-splicing of group II introns, proceed through the same two-step mechanism, involving the formation of a 2'-5' branched intermediate and product¹⁷⁻²⁰. In group II introns, all of the machinery necessary for splicing is contained within the intron, whereas pre-mRNA splicing takes place in spliceosomes, complex particles containing small nuclear ribonucleoproteins (snRNPs) and a large number of non-snRNP splicing factors^{21,22}. *Trans*-splicing seems to represent an intermediate evolutionary stage in which both the 5' exon and an snRNP-like region are contained within the SL RNA. If the functions of the U1 and U5 snRNPs are to recognize intron and exon sequences, respectively, at both the 5' and 3' splice sites²³⁻²⁵, then splice site recognition in *trans*-splicing may involve analogous activities within the SL RNA²⁵.

We have shown that accurately initiated and terminated SL RNAs are assembled into SL snRNPs in mammalian cells, and that a nematode 3' *trans* acceptor, as well as the normally *cis*-spliced adenovirus 3' splice site, are assembled into splicing complexes capable of functional interactions with SL RNAs. *In vitro* spliceosome assembly studies in mammalian cell extracts have shown that pre-spliceosome complexes containing specific splicing components can assemble on pre-mRNAs containing only a 5' or 3' splice site^{26,27}. In the *trans*-splicing reaction, complexes assembled on separate RNA molecules may interact by virtue of the utilization of novel splicing factors or unique interactions of known splicing components. For example, it has been shown that the nematode SL RNA interacts with the U6 snRNP (T. Nilsen, personal communication) in a region of U6 previously seen to base-pair with the U2 snRNP in mammalian splicing systems²⁸⁻³⁰.

In summary, we have shown that the mammalian splicing apparatus can carry out *trans*-splicing by a mechanism involving

SL RNAs. We are currently exploiting this observation to search for mammalian SL RNAs and *trans*-splicing acceptors. □

Received 11 September; accepted 12 November 1992.

- Huang, X.-Y. & Hirsh, D. *Genet. Engng.* **14**, 211-229 (1992).
- Agabian, N. *Cell* **61**, 1157-1160 (1990).
- Tessier, L.-H. *et al.* *EMBO J.* **10**, 2621-2625 (1991).
- Solnick, D. *Cell* **42**, 157-164 (1985).
- Konarska, M. M., Padgett, R. A. & Sharp, P. A. *Cell* **42**, 165-171 (1985).
- Solnick, D. *Cell* **44**, 211 (1986).
- Sharp, P. A. & Konarska, M. M. *Cell* **44**, 211 (1986).
- Shimizu, A., Nussenzweig, M. C., Mizuta, T.-R., Leder, P. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8020-8023 (1989).
- Sureau, A. & Perbal, B. C. *R. Acad. Sci. Paris* **312**, 323-328 (1991).
- Krause, M. & Hirsh, D. *Cell* **49**, 753-761 (1987).
- Milhausen, M., Nelson, R. G., Sather, S., Selkirk, M. & Agabian, N. *Cell* **38**, 721-729 (1984).
- Zhuang, Y. & Weiner, A. M. *Cell* **46**, 827-835 (1986).
- Sproat, B. S., Lamond, A. I., Beijer, B., Neuner, P. & Ryder, U. *Nucleic Acids Res.* **17**, 3373-3386 (1989).
- Lamond, A. I., Sproat, B., Ryder, U. & Hamm, J. *Cell* **58**, 383-390 (1989).
- Bruzik, J. P. & Steitz, J. A. *Cell* **62**, 889-899 (1990).
- Sharp, P. A. *Cell* **50**, 147-148 (1987).
- Cech, T. R. *Cell* **44**, 207-210 (1986).
- Murphy, W. J., Watkins, K. P. & Agabian, N. *Cell* **47**, 517-525 (1986).
- Sutton, R. E. & Boothroyd, J. C. *Cell* **47**, 527-535 (1986).
- Bektesh, S. L. & Hirsh, D. I. *Nucleic Acids Res.* **16**, 5692 (1988).
- Green, M. R. A. *Rev. Biochem.* **7**, 559-599 (1991).
- Reed, R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8031-8035 (1990).
- Reich, C. I., VanHoy, R. W., Porter, G. L. & Wise, J. A. *Cell* **69**, 1159-1169 (1992).
- Newman, A. J. & Norman, C. *Cell* **68**, 743-754 (1992).
- Steitz, J. A. *Science* **257**, 888-889 (1992).
- Fu, X.-D. & Maniatis, T. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1725-1729 (1992).
- Bennett, M., Michaud, S., Kingston, J. & Reed, R. *Genes Dev.* **6**, 1986-2000 (1992).
- Hausner, T.-P., Giglio, L. M. & Weiner, A. M. *Genes Dev.* **4**, 2146-2156 (1990).
- Wu, J. & Manley, J. L. *Nature* **352**, 818-821 (1991).
- Datta, B. & Weiner, A. M. *Nature* **352**, 821-824 (1991).
- Mangin, M., Ares, M. & Weiner, A. M. *EMBO J.* **5**, 987-995 (1986).
- Howe, J. G. & Shu, M.-D. *Cell* **57**, 825-834 (1989).
- Wasserman, D. & Steitz, J. A. *Molec. cell. Biol.* **11**, 3432-3445 (1991).
- Blencowe, B. J., Sproat, B. S., Ryder, U., Barabino, S. & Lamond, A. I. *Cell* **59**, 531-539 (1989).
- Bruzik, J. P., Van Doren, K., Hirsh, D. & Steitz, J. A. *Nature* **335**, 559-562 (1988).
- Hannon, G. J., Maroney, P. A., Denker, J. A. & Nilsen, T. W. *Cell* **61**, 1247-1255 (1990).

ACKNOWLEDGEMENTS. We thank J. Steitz for encouragement, technical advice and for reagents, R. Ach and A. Weiner for the pUC-U1 vector, and J. Steitz, X.-D. Fu and members of our laboratory for discussion and for comments on the manuscript. This research was supported by a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund for Medical Research (J.P.B.) and a grant from the NIH (T.M.).

CORRECTION

Homology of a 150K cytoplasmic dynein-associated polypeptide with the *Drosophila* gene *Glued*

Erika L. F. Holzbaur, James A. Hammarback, Bryce M. Paschal, Nancy G. Kravitz, K. Kevin Pfister & Richard B. Vallee

Nature **351**, 579-583

EXAMINATION of the sequence for a homologue of the polypeptide described in the above paper (EMBL accession number X62160) raised the possibility that a small segment of the rat sequence as originally reported was in an incorrect translational frame. We have closely examined the original data and conclude that two sequencing errors caused the reading frame for nucleotides 1,099-1,214 to be shifted by one base. The corrected form predicts an amino-acid sequence for residues 277-315 of EAK-EAKEALEAKERYMEEMADTADAIEMATLDKEMAEER.

The revised sequence now shows an even greater extent of homology with the predicted polypeptide product of the *Drosophila* gene *Glued*, as well as a greater extent of coiled-coil alpha helix consistent with the structure of the *Glued* gene product. Therefore none of the conclusions of the paper is altered. As a clearer designation for the rat polypeptide which we characterized, we now suggest the name p150^{Glued}. The revised sequence has been submitted to EMBL.

Culturing aids

Pre-coated coverslips, cell separation flasks, an assay for *in vitro* toxicity testing, microtomes, assorted media and media supplements — new products in the cell and tissue culture line.

ONCOGENE Science announces the introduction of SH2/SH3 BioTools, a product line that has been designed to help answer fundamental questions about cell signalling (*Reader Service No. 100*). SH2/SH3-containing proteins are critical components in the signal transduction pathway, serving as 'adaptors' between receptor tyrosine kinases and their substrates. They function by binding to specific phosphorylated sequences in the activated receptor tyrosine kinase molecule and mediating a specific signal response. SH2/SH3 domains have been identified in many cytoplasmic signalling proteins, as well as in oncoproteins such as *v-crk* and *vav*. The company says that with these products, researchers will be able to study directly the different binding specificities of proteins involved in signal transduction. Available next month, SH2/SH3 BioTools can be used to **detect, separate, purify and identify phosphoproteins** from activated cells. The line presently consists of two classes of products: SH2/SH3 agarose conjugates that may be used for immunoprecipitation or as affinity purification matrices for the identification and isolation of SH2 domain binding proteins from activated cell lysates. Biotinylated SH2/SH3 products can be used as detection systems for western blotting of cellular extracts.

The **osteocalcin assay** from Metra Biosystems is a new ELISA that measures bone turnover (*Reader Service No. 101*). Bone turnover has been shown to be important in osteoporosis, rheumatoid arthritis, osteoarthritis and Paget's disease. The company's osteocalcin assay measures *de novo* osteocalcin in serum or tissue culture media. The assay contains a conformation-dependent antibody to measure only newly synthesized osteocalcin, not osteocalcin released from resorbed bone. It is a horse-radish peroxidase-based competitive ELISA that uses a microtitre format.

Stem cell factor has been shown to play an essential role in gametogenesis, melanogenesis and early stages of haematopoiesis. It is also expressed in the nervous system, suggesting a possible role in the development of the nervous system. British Bio-technology is now marketing a **kit for labelling and studying stem cell factor receptors** on cell surfaces by flow cytometry (*Reader Service No. 102*). The Fluorokine kit can be used to determine the percentage of cells bearing cytokine receptors within a population and also the receptors density for human stem cell factor on cell surfaces. The kit contains biotinylated stem cell fac-

tor, detection reagents (avidin fluorescein) and wash buffers that are optimized to lower nonspecific binding and to ensure the sensitivity necessary to detect low numbers of cytokine receptors. The labelled cytokine then binds to the cells via specific cell surface receptors with high specificity such that fluorescent intensity is proportional to the density of the receptors. BBt also supplies the recombinant protein; both murine and human sequences are available, with or without carrier protein. Both proteins are produced in *Escherichia coli* and purified to greater than 97 per cent purity by silver-stained SDS-PAGE. The recently launched polyclonal antibody to human stem cell factor is raised in goats and the immune sera purified by Protein-G affinity chromatography.

■ Cell culture/separation

Collaborative Biomedical Products/Becton Dickinson now offer Biocoat Brand **coverslips coated with rat-tail collagen type I or with human fibronectin** (*Reader Service No. 103*). These coverslips are designed to provide the convenience and consistency of extracellular matrix-coated cell



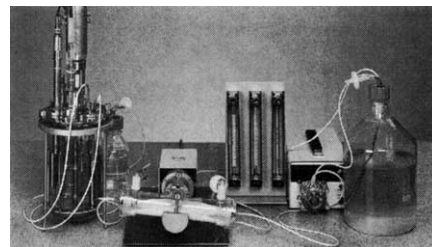
Pre-coated glass coverslips.

culture surfaces for studies of cellular attachment, spreading, growth and/or differentiation, cell-extracellular matrix interactions, receptor-ligand binding assays, drug screening and sensitivity assays, and primary tumour culture. Extracellular matrix-coated coverslips may also be useful in facilitating the growth of amniocytes. Each lot of Biocoat ECM-coated coverslips is tested for its ability to promote attachment and spreading of HT-1080 cells (collagen I) or BHK-21 cells (human fibronectin).

Applied Immune Sciences is adding a T-150 size to its line-up of MicroCELLector **cell separation flasks** that are available to researchers (*Reader Service No. 104*). The company sells these sterile flasks with various purified monoclonal antibodies or other ligands that bind human cells or other proteins covalently bound to interior surfaces.

Flasks are also available with the proprietary activated surface for use by researchers with their own antibodies or ligands. As the antibodies or ligands are permanently bound to the polystyrene flask, adherent cells are negative for surface monoclonal antibody after release. AIS says that the larger surface area of the new flasks will reduce the quantity of flasks, and processing time, for researchers who need large cell populations. The T-150 flasks are available with anti-CD4 or anti-CD8 antibodies, or with soybean agglutinin (SBA). The company now offers a stem cell introductory kit containing four SBA and two CD34 T-25 flasks.

Applikon has introduced a **multiple-fibre bioreactor** that has been specially designed as a cell culture system for the production of monoclonal antibodies and other mammalian cell products (*Reader Service No. 105*). The Applikon MFB features extra capillary volume, a long-life cartridge and direct aeration to avoid anoxic pockets. According to the company, the MFB can produce up to 58.6 mg of monoclonal antibodies per day. This rate of production is 20-times higher than that obtained with a stirrer flask, Applikon says. Unlike other cartridge systems, the MFB contains three bundles of fibres: one for the rapid transfer of nutrients, the other two are for aeration. All bundles can be operated independently. A gentle mixing action is designed to provide homogeneous conditions. The bioreactor can be operated either



High-yielding multiple fibre bioreactor.

in a circulation mode or in a single-pass mode. With the former mode, Applikon says that a high production efficiency can be achieved with minimum volumes of media. In the single-pass mode, only fresh medium enters the bioreactor at a very low rate, thus maximizing the nutrient transfer rate. The MFB can be autoclaved.

■ Toxicity testing

The NeutralRed Bioassay for *in vitro* toxicity testing from Clonetics is now available with pre-seeded 96-well plates of normal

human cells (*Reader Service No. 106*). Clonetics says that the new pre-seeded version has advantages over the original, where the cells are shipped in T-25 flasks. It offers a 50 per cent reduction in assay time, enhanced accuracy and economy and reduced potential for error. The new format eliminates the need to subculture or transfer the cells. The pre-seeded plates can be used directly in the assay from start to finish and the assay results are read directly in the same plates. Clonetics says that this saves time previously required for transferring the cells from the T-25 flask into the 96-well plates and eliminates the possibility of destroying cells through over application of trypsin (the enzyme used to detach the cells from the plastic T-25 flask).

Promega announces the addition of the CellTiter 96 AQ_{eu} **nonradioactive cell proliferation assay** to its line of cellular regulation products (*Reader Service No. 107*). The assay, which is based on the use of a novel tetrazolium, MTS, is a colorimetric method for determining the number of viable cells in proliferation or chemosensitivity assays. Promega says that in addition to the advantages over competing technologies of nonradioactivity, speed, safety and convenience, the assay provides flexibility in that plates can be read and returned to the incubator for further colour development. Viable cells can be recovered after the assay.

■ At the cutting edge

Carl Zeiss has brought out a line of MICROM **electronic microtomes** that are designed to improve the productivity of laboratories engaged in sectioning specimens (*Reader Service No. 108*). Two models are available: the HM-340E multipurpose rotary microtome for routine and research applications and the HM440E



Electronic microtomes by Carl Zeiss.

sliding microtome. The HM340E features a removable touchpad operating panel, a large, removable, integrated section waste tray, automatic approach system of the specimen from any position, pre-selected coaxial fine/trim feed and a removable storage unit with a cooled surface. The HM440E features motorized vertical specimen movement, specimen retraction in the sledges return mode, automatic section thickness feed, one button for up/down movement, fine feed

and trimming, pre-selected coaxial fine/trim feed and an integrated section waste tray.

The new **oscillating tissue slicer** from Electron Microscopy Sciences, which is supplied with or without refrigeration, is designed to offer a cost-effective solution to sectioning fresh or fixed tissue (20 µm for fresh tissue), without the need for embedding or freezing (*Reader Service No. 109*). For maximum precision, EMS says that its tissue slicer features thumbwheel selection of section thickness, including motorized control of the blade height and digital display of the total cutting excursion. All controls can be set with one hand; the blade oscillation is activated with a foot switch, freeing the operator's other hand to manipulate the specimen. Operators have the option of either mounting the specimen directly to a three-level step tissue pedestal or attaching the sample vice holder and adhering the specimen directly to an anodized aluminium block. The instrument includes a magnifier and a halogen lamp. The thermoelectric refrigeration system maintains the temperature of the tray near 0 °C during sectioning.

■ Media and media supplements

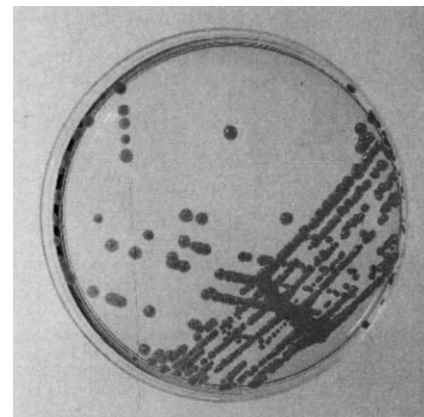
Biochrom has developed a **tissue culture medium with a stable glutamine source**: an acetylated dipeptide replaces the usual L-glutamine (*Reader Service No. 110*). The company says that the dipeptide N-acetyl-L-alanyl-L-glutamine is stable even at high temperatures. The liquid medium with stable glutamine compared favourably when tested against medium with glutamine provided in the usual manner in cell growth tests using hamster kidney cells (BHK 21), RITA cells (monkey kidney epithelial cells) and hybridoma cells (3C3).

BriClone is a **hybridoma cloning medium** supplied by TCS Biologicals that is designed to reduce the time and effort required to clone cell lines (*Reader Service No. 111*). It eliminates the need to prepare feeder cells for use in the production and cloning of hybridomas. TCS Biologicals says that BriClone provides increased efficiency in cloning by limiting dilution methods and its use minimizes the risk of culture contamination. Consistent results are achieved as there is no batch variation and no depletion of nutrients as can occur with the use of feeder cells.

Dr R. Pellkofer has developed a complete **serum-free medium**, medium of defined components (MDC), which shows similar growth-promoting activities as conventional media containing fetal calf serum (*Reader Service No. 112*). The company says that MDC offers an improved formulation over that developed by Iscove and Melchers, where the lipids are added in the form of a cloudy suspension of liposomes.

By keeping the lipids in a clear and stable form, Pellkofer says he has managed to produce a serum-free medium that offers not only significantly improved growth performance of the cells but also a longer shelf life. Three variants are offered and all are supplied in a ready-to-use form where there is no need for further supplementation.

Unipath had added **yeast and mould agar** to its Oxoid range of culture media (*Reader Service No. 113*). Yeast and mould agar is based on the formulation described by Wickerham^{1,2} and is recommended for the isolation and maintenance of yeasts and



Typical colonies of *Aspergillus niger* and *Saccharomyces cerevisiae* growing on yeast and mould agar.

moulds. The medium has a pH of 6.2 ± 0.2 and can be rendered selective by the addition of acid to reduce the pH to 4.0. A suitable acid for use is Oxoid sterile lactic acid.

1. Wickerham, L. J. *US Dept. Agric. Tech. Bull.* No. 1029, 1-19 (1951).
2. Wickerham, L. J. & Rettger, L. F. *J. Trop. med. Hyg.* **42**, 174-179 (1939).

Northumbria Biologicals offers two **supplements for producing serum-free media** for hybridoma cell cultures (*Reader Service No. 114*). Originally developed for serum-free culture of helper T-cells, SF-1 serum-free medium can be used for hybridoma cell proliferation and monoclonal antibody production. SF-X protein-free supplement is both serum-free and protein-free: the purification of monoclonal antibodies is simplified as contamination by proteins, or factors bound to serum proteins, is completely eliminated. SF-X is added to Iscove's/F-12 (1:1) medium (also protein-free) to produce a medium that is capable of supporting the growth of hybridoma cell lines. SF-1 is synthetic and the constituents are completely diffused. The company says that batch-to-batch variation is minimal, the protein content is less than 2.6 mg ml^{-1} and there are no immunoglobulins present.

LipoMAX bovine lipoprotein supplement from Gibco BRL is a source and delivery vehicle of cholesterol in serum-

ADVERTISEMENTS

**\$3.50/Base
Custom DNA
Synthesis**

No Set-up Fee,
Shipped in 24 hours.

1-800-788-4362
1-800-786-4362 Fax

KEYSTONE
Laboratories, Inc.

Reader Service No.33

Custom made
PEPTIDES

Around the globe
\$30.00 /residue
Shipped in 10 days

- 8 to 20 amino acids
- 20 µmole scale
- HPLC analysis & assured quality
- Since 1984

Your Source of Peptides
BIO-SYNTHESIS
1-800-227-0627 Fax 214-429-0442

Reader Service No.52

Zeineh 10/20/DNA/TLC
HAND-HELD SCANNING DENSITOMETER

To quantify:
Gels, blots, TLC, X-Ray film, Photo.

1-D pkg price \$995

8 bits, 60 um resolution, 0-3 OD, sample 11x40 cm, IBM or clones
Options: UV, 2-D, DNA, ELISA, colony counting..... \$995 each.
Universal software: for any scanner. Main 1D, 2D, DNA, \$995 each.
Representatives needed, 30% discount of USA price list.

Do-it-all : Departmental, Laser + Video
Micro Densitometry Station
\$ 16,990

Hi-Speed
Soft Laser

imaging
Video

BIOMED INSTRUMENTS INC. 1020 S Raymond Av, #B, Fullerton
CA 92631. Pnn: (800) 927-2290, (714) 870-0290, Fx: 17148706385

Reader Service No.22

**Highest Quality
Oligonucleotides**

Synthesised, purified and shipped in 3 days

£2.75 per base
(All inclusive)

0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit
King's College School of Medicine & Dentistry
Tel. 071 - 326 3126 FAX 071 - 733 3877

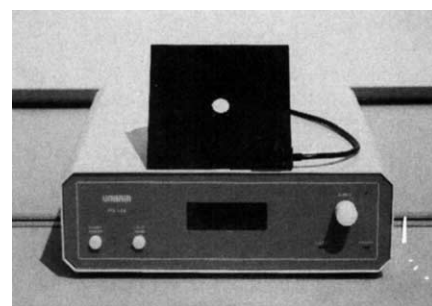
Reader Service No.56

free media (*Reader Service No. 115*). It contains both high-density (HDL) and low-density lipoprotein (LDL) components, cholesterol, and lipid moieties purified from bovine plasma. It is heat-treated at 60 °C for 10 h, then filtered three times through 0.1-µm membrane filters and aseptically dispensed into PETG bottles. LipoMAX supplement is concentrated to not less than 20 g l⁻¹ and is typically diluted for use ~400 to 2,000-fold to yield a final concentration 10–50 µg ml⁻¹ protein. Gibco BRL says that LipoMAX should play a valuable role in the regulation and maintenance of cellular growth and metabolism. As it has both HDL and LDL fractions, LipoMAX helps in the proliferation and growth of many cell types in serum-free or serum-reduced media. The cell types include suspension, anchorage dependent, primary normal and transformed cell lines.

Intergen announces the introduction of **iron-enriched calf serum**, which is designed to provide a media supplement that does not compromise critical growth promotion characteristics (*Reader Service No. 116*). The iron-enriched serum is collected from formula fed calves, which increases the transferrin levels in their blood and allows for higher iron binding than typical fetal or calf serum. Intergen says that the serum has been shown to be similar or equivalent to fetal bovine serum, using its panel of indicator cell lines.

■ **Equipment**

The Linkam TH60 **precision warm stages** can be used with all optical microscopes and are offered in a range of sizes with the new MS100 controller (*Reader Service No. 117*). The warm stages have a temperature range from ambient to 60 °C (100 °C on request), are 1.4-mm thick and are fitted with 0.5-mm thick feet to prevent the microscope table from being heated. The bifilar heating element embedded within the stage helps prevent heater-induced currents from affecting the sample, while the construction of the heating element ensures an even heat distribution over the surface area of the stage. The MS100 controller will maintain the



Linkam's warm microscope stage.

temperature of the stage to better than 0.1 °C over the entire temperature range. It displays the temperature or limit value to 0.1 °C. Warm stages that fit directly into Nikon and Olympus inverted microscope tables are also available.

ICN Flow has extended its line of **laminar safety cabinets** with the launch of the new Class 100 model for tissue culture applications (*Reader Service No. 118*). Called the Gelaire TC and built on the same compact chassis as the other ICN Flow cabinets, the new model is fully upgradable to a BSB cabinet. The work surface is perforated stainless steel to catch spills. A four-bulb illumination system is designed to enhance the natural light provided by having glass side panels. A switchable mains socket is provided within the cabinet. All controls are mounted on a soft touch-control panel with a surface that can be wiped clean. The unit can be supplied with or without an exhaust HEPA filter and does not require ducting to an outside wall.

■ **Assorted reagents**

Histoclear II is a new nontoxic, nonflammable and biodegradable **histological clearing agent** from National Diagnostics (*Reader Service No. 119*). Whereas xylene may leave tissue hard and brittle, the company says that Histoclear II leaves tissue flexible enough for cutting into thin sections. Histoclear II is miscible in all proportions with ethanol, isopropanol and t-butanol. It is also miscible with all paraffin-based tissue-embedding media and all permanent mounting materials.

Remove even the most stubborn of stains, including acidic stains and Schiff's reagent, from skin and clothing with the new range of **stain removers** from Shandon Scientific (*Reader Service No. 120*). The stain remover for basic stains such as haematoxylin, crystal violet and blood stains, is a cream that decolourizes and removes the dye molecule rapidly. For acid stains such as eosin, trichromes and Papanicolaou, which penetrate the skin, a new cream rubbed onto the stained area releases the dye. Completing the line-up is a dry, granular stain remover for removing Schiff's reagent. □

For more information, fill in the reader service card bound inside the journal.

Oligos Etc. Inc.

DNA \$4.50/base
S-DNA \$8/base
Methyl Phosphonate DNA \$8/base
RNA \$8/base
S-RNA \$12/base
O-Methyl RNA \$15/base
Chimerics, Biotin
Modifiers, Fluorescent Tags etc .
No Set-Up Charges

800-888-2358/fax 800-869-0813

Intl.fax 203-438-3637

Reader Service No.30